Analysis of the Large Gamma Ray Flares of Mkn 421 as Observed with HEGRA CT1 on the Island La Palma in 2001

Ph.D. Thesis

by

Thomas Schweizer

supervised by

Manel Martinez
IFAE
Edifici Cn., UAB
08193 Bellaterra/Spain

and

Eckart Lorenz
Max Planck Institut für Physik
80805 München/Germany

Acknoledgements

First, I wish to sincerely thank my supervisors Dr. Manel Martinez from the Institut de Fisica d'Altes Energies/Spain and Dr. Eckart Lorenz from the Max Planck Institut für Physik in Munich for supervising and providing me with this interesting Ph.D. subject and for giving me a chance to do real physics for my doctoral work. Both helped me immensely by providing me with the necessary work environment and giving me the possibility to participate in several conferences and schools. In addition, they always helped me with questions and problems related to physics and last but not least for cross-reading this thesis. Dr. Eckart Lorenz provided me with an abundance of long distance support via many fruitful discussions and invitations to Munich. Without his help it would have been much more difficult to accomplish good results. Many of the ideas in this work, like weights in the calculation of Hillas parameters, and several more came from him.

I also wish to thank Enrique Fernandez very much for providing me with the opportunity to do a doctoral thesis at IFAE, to live in Spain and to learn about and experience two fascinating cultures, the Catalan one and the Spanish one.

In addition I would like to thank all the members of the HEGRA collaboration for supplying me with the astronomical data. I am grateful to all the HEGRA members of the Max Planck Institut für Physik in München that I have known personally, for very good cooperation and for teaching me an abundance of valuable information about Cherenkov Telescopes, especially the CT1. These people include Razmik Mirzoyan, Martin Kestel, Daniel Kranich and Juan Cortina. I owe special thanks to Wolfgang Wittek for some very helpful discussions and for cross-reading my thesis.

At this point I wish to sincerely thank Kristo Karr for the painful English correction of my thesis. Without him the thesis would not be what it is now! The other cross readers must also to be mentioned here. These people are Markus Gaug and Martin Kestel.

At IFAE I wish to thank Markus Gaug, David Paneque, Josep Flix, Javier Lopez, Eva Domingo and Oscar Blanch for their good team work and for providing a very nice working environment at the institute. In this context I owe special thanks to Sacha Ostankov with whom I had a very pleasant and productive working collaboration.

Finally, I wish to extend my most heartfelt thanks to my wife Miha who supported me with her love, which provided me with the energy needed to accomplish this work. Another heartfelt thanks to my parents for supporting my studies in physics, helping me during difficult times and for providing steadfast moral support.

Contents

1 Introduction to the theory and detectors **11**
- 1.1 Cosmic rays . . . 11
  - 1.1.1 The spectrum of cosmic rays . . . 12
  - 1.1.2 Sources of cosmic rays . . . 12
  - 1.1.3 γ-rays and γ-ray production processes . . . 13
- 1.2 Active galactic nuclei and the class of BL LACs . . . 14
  - 1.2.1 Classification of active galactic nuclei (AGN) . . . 14
  - 1.2.2 Disc dynamics and the expulsion of jets . . . 15
  - 1.2.3 Jet models and the SSC model . . . 17
  - 1.2.4 Jets of Mkn 421 and Mkn 501 as seen by microwave telescopes . . . 21
  - 1.2.5 Fast variablity and other properties of AGN flares . . . 22
- 1.3 The Gamma Ray Horizon . . . 24
- 1.4 Theory of air showers . . . 29
  - 1.4.1 Electromagnetic cascades . . . 30
  - 1.4.2 Hadronic cascades . . . 34
  - 1.4.3 Cherenkov light production . . . 36
- .1 The principle of Cherenkov imaging telescopes . . . 36
- .2 The HEGRA experiment: An overview . . . 38
- .3 The HEGRA CT1 telescope . . . 40
  - .3.1 Technical details . . . 40
  - .3.2 The trigger of CT1 . . . 41
- .4 The All Sky monitor of the RXTE satellite . . . 41

A The Analysis of shower images **43**
- A.1 The calibration of the telescope data . . . 44
- A.2 Monte Carlo Simulation of air showers . . . 44
  - A.2.1 Simulation of the night sky light . . . 45
  - A.2.2 Trigger efficiencies, cut efficiencies and effective areas . . . 46
- A.3 Gamma/hadron separation methods . . . 48
  - A.3.1 The classical Hillas parameters to describe shower images . . . 48
  - A.3.2 Image cleaning algorithms to remove the night sky background in the camera . . . 53
  - A.3.3 Static cuts and dynamical cuts: Dependence of the Hillas parameters on the energy, the zenith angle, the impact parameter and the night sky background . . . 54
  - A.3.4 Background determination by using the ALPHA-plot . . . 58
  - A.3.5 The linear discriminant analysis as a dynamical cut and as a tool to quantify the discrimination power of image parameters . . . 60
  - A.3.6 Comparison of the discrimination power of the LDA and the dynamical cuts, tested with real Mkn 421 data . . . 64
  - A.3.7 Conclusion about the efficiency of the LDA separation method and an outlook . . . 69

A.4 New image parameters and cleaning algorithms . . . 71
A.4.1 Introduction of weights in the calculation of Hillas parameters . . . 71
A.4.2 The head-tail asymmetry defined in a such way that it overcomes the finite resolution of the CT1 camera . . . 80
A.4.3 The problem of truncated images due to the limited size of the CT1 camera and the leakage parameter . . . 81
A.4.4 Mountains and islands in the shower image: Difference between electromagnetic showers and hadronic showers . . . 83
A.4.5 Improvement in gamma/hadron separation . . . 88
A.4.6 Conclusion about the introduction of new image parameters and new image cleaning algorithms . . . 91
A.5 Energy estimation by the least squares method . . . 94
A.5.1 Conclusion about the energy estimation using the linear least squares method . . . 95
A.6 Mispointing of the telescope and its correction . . . 95
A.7 Differential flux spectrum calculation . . . 101
A.7.1 Determination of the energy excess event distribution . . . 101
A.7.2 Unfolding the spectrum . . . 102
A.7.3 The spectrum, fitting and reverse check of the result . . . 105
A.7.4 Discussion of systematic errors and the reliability of the obtained spectrum . . . 107
A.8 Calculation of lightcurve and hardness ratio . . . 108
A.8.1 The mathematical background of integrated flux measurements . . . 108
A.8.2 Discussion about systematic errors and the reliability of the measured lightcurve . . . 111

B Analysis of the flares of Mkn 421 in 2001 **113**
B.1 Light curve and timing of the flares of Mkn 421 . . . 113
B.1.1 The light curve of Mkn 421 as measured by CT1 . . . 113
B.1.2 Comparison of the lightcurve to the one measured by the CT-system 122
B.1.3 The 2 keV-12 keV lightcurve measured by the All Sky Monitor of the RXTE satellite . . . 122
B.1.4 The discrete correlation function applied to TeV and X-Ray light curves . . . 124
B.1.5 Conclusion about the γ-ray/x-ray correlation studies . . . 133
B.1.6 The hardness ratio as measured by CT1 in 2001 . . . 134
B.2 Spectrum of Mkn 421 . . . 138
B.2.1 The spectrum averaged over all flares . . . 138
B.2.2 The Crab nebula spectrum for consistency cross check . . . 141
B.2.3 Analysis of the Mkn 421 spectrum during different flare states . . . 141
B.2.4 Conclusions and a discussion of the results . . . 147

C Comparison with Mkn 501 and Conclusion **149**
C.1 Spectral properties of γ-flares . . . 149
C.1.1 Time scales, correlations and jet models . . . 153
C.2 Outlook . . . 157

D Summary **159**

E Appendix A: Theory and calculations **163**
E.1 Verification of the NSB distribution function . . . 163
E.2 The synchrotron self Compton (SSC) model . . . 165
E.3 Error calculation of the correlation function . . . 167
E.4 Re-binning and averaging of flux bins . . . 168

F Appendix B: Lightcurves and ALPHA plots **169**
F.1 The complete daily lightcurve 170
F.2 The estimation of the background for the spectrum 182

Abbreviations and shortcuts

ACT	Air Cherenkov telescope
ADC	Analog to digital converter
AGN	Active galactic nucleus
ASM	All sky monitor of the RXTE satellite
CIB	Cosmic infrared background
CMB	Cosmic microwave background
DFT	Discrete Fourier transform
EAS	Extended air shower
EGRET	Energetic Gamma-ray Experiment Telescope
FOV	Field of view
FSRQ	Flat spectrum radio quasar
FWHM	Full width half maximum
GRB	Gamma ray burst
HEGRA	High energy gamma ray astronomy
HV	High voltage
NSB	Night sky background
LDA	Linear discriminant analysis
MC	Monte Carlo simulation
NKG	Nishimura Kamata Greisen
PhE	Photoelectron
QE	Quantum efficiency
SNR	Super nova remnant

Chapter 1

Introduction to the theory and detectors

In recent years Astroparticle Physics has been a field of steady growth and increasing interest. Only since the last 15 years, has one been able to measure and observe phenomena which are completely new and which have significant impact to our picture of the universe, distant galaxies and black holes. In my thesis I will report on the analysis of the large, persistent very high energy gamma-ray (shortcut γ-ray) flaring of the active galactic nucleus ***Mkn 421.***

In this chapter I wish to describe the essential theory which is needed for this Thesis. First the properties and the origin of cosmic rays will be summarized. This is followed by a classification and description of the active galactic nuclei (AGNs). A very simple jet model which accelerates the parent particles of very high energy (VHE) γ's[1] will be described. This is needed to understand the properties of Mkn 421. The very high energy gamma ray absorption by the cosmic infrared background, which might affect the observed spectrum, is quantified and simulated. Then I will explain the various possibilities for detecting cosmic rays and describe the HEGRA experiment. Later in this chapter, Cherenkov telescopes will be discussed, including how they function. Emphasis is put on the HEGRA CT1 telescope which was used to record the data used in this thesis.

1.1 Cosmic rays

Every second about 1000 high energy cosmic particles (integrates flux above 10 GeV) per square meter hit the Earth's atmosphere. Up to an energy of approximately 1 PeV the chemical composition has been measured directly (for an overview see [LON97/1]). Most of them are ionized nuclei. Approximately 90% of them are protons, 9% are alpha particles, a small fraction are ionized nuclei of heavier elements and a fraction of a percent (0.1%) are high energy gammas [Gai90]. Neutrinos (ν's) are also present but their number is unknown because of their extremely small interaction cross section.

Charged particles are deflected in the intergalactic and galactic magnetic field. Therefore, they no longer point back to their origin. Only neutral particles can be extrapolated back to their sources. Among the neutral particles, neutrons[2], ν's and γ's, only the latter type are sufficiently abundant. γ's are ***messenger particles*** which give us information about the source, its flux, spectrum, timing information and about the γ propagation in the path between the source and the observer.

[1] Throughout this document the abbreviation 'γ' refers to a high energy (> 1 GeV) photon.

[2] Neutrons decay with a life time of 940 s. Still, extremely high energy neutrons (>10^{18} eV) could reach us before they decay from close objects as the center of our galaxy.

The main questions today are: (1) Where do they ***come*** from ? (2) What is the ***composition*** of cosmic rays ? and (3) What are the ***acceleration mechanisms*** that are able to accelerate particles to these high energies ?

1.1.1 The spectrum of cosmic rays

As of today, it is ***not fully known*** where the cosmic rays (CR) come from and how they are accelerated to such high energies (in the GeV to PeV energy range). It is commonly agreed that there exist three populations of cosmic rays: (1) Particles emitted by the sun which relatively low energies up to about 10 GeV, (2) galactic cosmic rays which dominate the spectrum up to 1 PeV and (3) extragalactic cosmic rays which have highest energies [LON97/2]. For simplicity I will restrict myself to the latter two types above energies of 10 GeV which are relevant for Cherenkov telescopes. Lets have a closer look at the spectrum of cosmic rays.

Fig. 1.1 shows the all particle spectrum of cosmic rays. The typical spectrum follows a power law and it is usually characterized by a constant called spectral index α:

$$\frac{dN}{dE} \propto E^{-\alpha} \tag{1.1}$$

Of interest is the region around 5 PeV, the so-called knee position, where the spectral index changes from $\alpha \approx 2.7$ to $\alpha \approx 3.0$. The part ***below*** the knee is believed to be quite well understood [LON97/1]. The dominant portion of the particles is of ***galactic origin***. The current belief is that shock wave acceleration in supernovae remnants (SNR) is the major source of galactic CR. Particles and ionized nuclei are emanated by red and blue giants and are boosted to higher energies in the shock fronts of SNRs. The chemical composition of the CR up to the knee can be explained by a combination of intrinsic source composition and a ***spallation process*** whereby heavier elements are broken down to lighter elements. It is believed that the CR are deflected and ***confined*** in the magnetic field of our galaxy. By measuring the abundance of radioactive isotopes it has been found that CR are confined from 10^6 to 10^7 years inside our galaxy. This explains the high isotropy of the CR in this energy range.

Above 1 PeV, shock wave acceleration becomes ineffective and the CR ***leak out*** of our galaxy (leaky box model [LON97/1]) because the galactic magnetic field is not strong enough to confine the particles in our galaxy. The chemical composition has not been measured above the knee and it is not clear as of today where these cosmic rays come from. The ***spectral index*** of the cosmic rays ***above the knee*** can be explained by the assumption that very high energy cosmic rays escape from the galaxy. The strongest argument for this theory is the ***increasing anisotropy*** with increasing energy of the particles [LON97/1]. A large part of the CR up to 10^3 PeV would therefore still be of galactic origin.

Above 10^4 PeV it is widely agreed that these CR must be ***extragalactic*** since their Larmor radii in the galactic magnet field is larger than the size of our galaxy [Gai90] and can ***no longer be confined***. Another argument is that the ***anisotropy changes*** and that the highest energy particles appear to originate mainly ***from high galactic latitudes***, most likely from the ***local super-cluster*** [LON97/1], while for lower energies the anisotropy points to the ***galactic plane***. For the highest energy particles (up to 10^{21} eV) no satisfactory acceleration model exists.

1.1.2 Sources of cosmic rays

Candidates for cosmic accelerators are objects that have significant magnetic fields and/or very large extensions and at the same time have extremely fast moving shock fronts. The most important candidates for the production of high energy CR are the following.

- ***Supernova remnant***s which accelerate particles in their shock front.

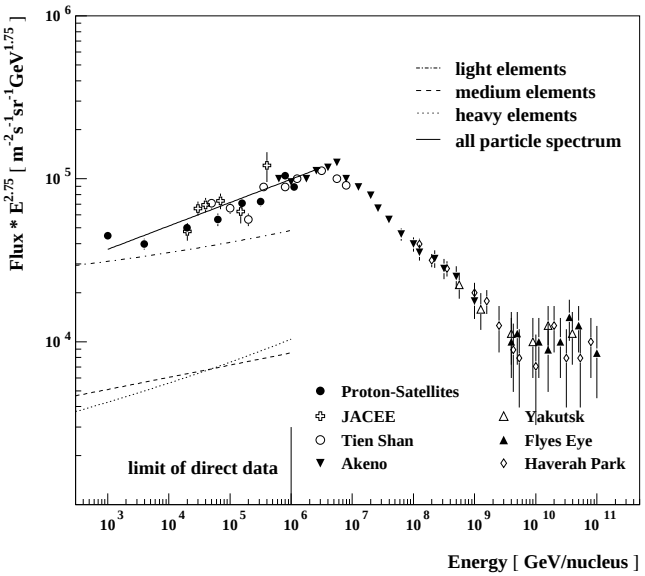


Figure 1.1: All particle differential spectrum multiplied with $E^{2.75}$ for better illustration of the knee at approximately 5 PeV (Taken from [Wie94]).

- ***Pulsars*** and neutron stars.
- ***Active galactic nuclei (AGN):*** Probably super massive black holes with more than 10^6 solar masses that emit plasma jets with high Lorentz factors (up to $\sim$10). Shock fronts inside these jets are able to accelerate particles to very high energies.
- ***Gamma ray bursts:*** Huge explosions of still unknown origin, visible almost uniquely in the soft γ-regime.
- ***Binary*** star systems ***with a neutron star*** or a ***black hole***.

In this work I will discuss active galactic nuclei (AGN), with focus on Mkn 421 which belongs to the group of BL LAC objects, for which the jet is pointing towards the earth.

1.1.3 γ-rays and γ-ray production processes

As already mentioned, the only particles which are sufficiently abundant and which can be traced back to their origin are γ-rays. Therefore γ-spectroscopy is the method to learn more about the physics of cosmic sources. There are several physical processes which are believed to generate gamma ray photons. These are:

- π^0 ***-decay:*** High energy protons interact with matter and produce a variety of hadronic particles. Approximately 30% of them are π^0, which decay almost instantaneously into two gammas (with branching ratio > 99%).
- ***Bremsstrahlung:*** If a charged particle is accelerated or decelerated (in an electric field), it emits photons. This could be a high energy electron or proton in the Coulomb field of a nucleus or ion. The of Bremsstrahlung spectrum of high energy electrons has the same spectral index as the electron spectrum itself, provided that the latter follows a power law.

- ***Synchrotron radiation:*** This is similar to the acceleration in an electric field: If a charged particle is accelerated (deflected) in a magnetic field, it emits photons. The peak emission of synchrotron radiation is given by:

$$E_{Peak} = 5 \cdot 10^{-9} \cdot H \cdot \gamma^2 \tag{1.2}$$

  H is the perpendicular magnetic field component in Gauss, γ is the γ-factor of the electron and the energy E_{Peak} is given in eV. For example, in the relativistic jet of an AGN, the magnetic field is about H=0.1 G. The electrons are accelerated by shock-waves up to $\gamma = 10^6$. The peak emission of synchrotron radiation occurs at approximately $E_{Peak} = 500\,eV$. One can see that in order to get high energy photons (> 1 GeV), relativistic electrons and a reasonably strong magnetic field are required. This is normally not the case. On the other hand, these low energy photons may produce a dense photon field with high energy electrons that they become a target for interaction: Namely, for ***inverse Compton scattering***. This process becomes very important for the generation of ultra high energy gamma rays.

- ***Inverse Compton scattering:*** Relativistic electrons up-scatter low energy photons to higher energies. The synchrotron radiation and the inverse Compton scattering will be discussed in detail in the section ***concerning*** the synchrotron-self-Compton (SSC model).

All these processes can take place in active galactic nuclei, which will be described in the following section.

1.2 Active galactic nuclei and the class of BL LACs

In this section I will briefly summarize some relevant features of ***active galactic nuclei (AGN)***. The term AGN refers to a class of galaxies which have a ***very large mass*** in the order of 10^6 to 10^{10} solar masses concentrated in a region of the ***size of our solar system***. It is believed that these central objects must be super massive black holes. The energy source of these objects is thought to be converted gravitational potential. There are several different classes and types of AGN which are all believed to be connected by a ***single model,*** namely the unified AGN scheme [Pad1/99, Pad2/99].

1.2.1 Classification of active galactic nuclei (AGN)

The classification [Pad1/99, Pad2/99, Gui98] is based on the galaxy shape and various physical properties like emission lines, microwave emission spectra and gamma ray emission (See fig. 1.2)

All of these objects are divided in two large groups: Spiral galaxies and elliptical galaxies. Members of the first group are called Seyfert galaxies and these are subdivided into Seyfert galaxies I and II. Type I galaxies have broad emission lines while the latter type galaxies have narrow emission lines. The elliptical galaxies are subdivided, according to their radio emission, in weak and strong radio emitters. Members of the class with weak radio emission are called Radio Quiet Quasars. Members of the class with strong radio emission are again subdivided in two big groups, one with strong and one with weak optical emission lines. The ones with strong optical emission lines are called Radio Quasars, which themselves are subdivided in two groups: ***Flat Spectrum Radio Quasars (FSRQ)*** and Steep Spectrum Radio Quasars (SSRQ). The first type, having weak or no optical emission lines are grouped into Radio Galaxies having no gamma emission, while the ***BL LAC*** objects which have some gamma emission, into a group with flat radio spectrum and optical polarization. Finally, BL LACs and FSRQ together form the class of ***Blazars.*** All blazars emit γ-rays.

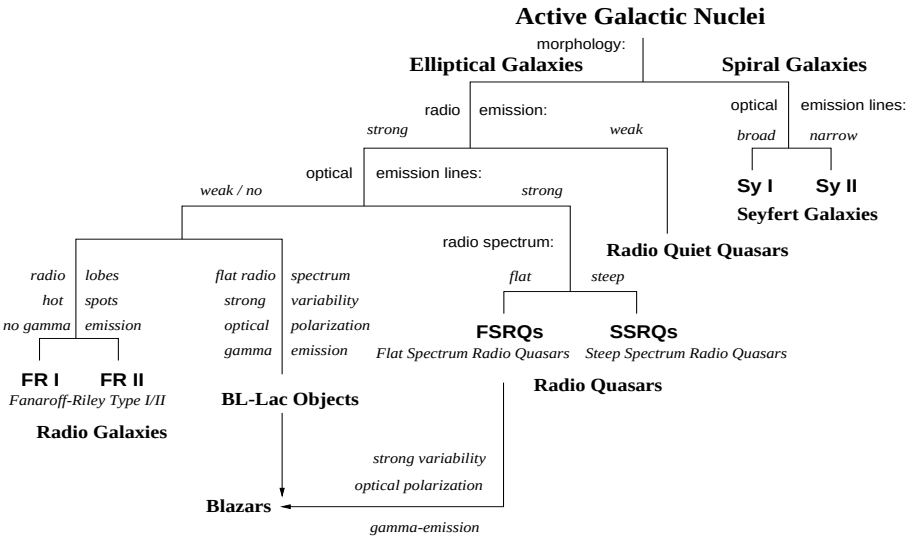


Figure 1.2: Classification of AGN, taken from [Böt97]

It has been tried to relate all these different classes and subdivisions to a general picture of AGNs. Fig. 1.3 shows a schematic sketch. At the center of each AGN is a ***super-massive black*** hole between 10^6 to 10^{10} solar masses surrounded by a ***highly relativistic rotating accretion disc*** [Liv02]. Due to ***friction*** the disc heats up to temperatures in the order of keV, thus transforming the gravitational potential into thermal energy. The plasma emits a ***thermal spectrum*** which peaks in the (soft) ***x-ray*** region. Photons emerging from the disc ***excite*** atoms and molecules in gas clouds, either close to the disc, which yields strong Doppler broadened optical emission lines, or further away from the torus, which produce narrow lines (because they are colder and have smaller velocities).

The different spectral differences are believed to originate from ***different observation angles*** and also from ***different physical parameters***, like the ***accretion rate*** and the ***spin*** of the black hole.

Sometimes there are strongly collimated, ***highly relativistic jets*** or blobs emitted perpendicular to the accretion disc from the poles of the black hole. The ***radio emission is related to synchrotron emission*** of relativistic (i.e. electrons) particles in the jet. The γ-emission is also believed to be produced by the jet. Since the jet has relativistic velocities, the radiation is beamed in forward direction and the γ-emission can only been seen if the observation angle to the jet is less than $\theta \sim \frac{1}{\Gamma}$(-> Blazars), where $\Gamma$is the Lorentz factor of the jet or blob.

The acceleration mechanism is ***yet not fully understood.*** The general ideas are presented in the next section.

1.2.2 Disc dynamics and the expulsion of jets

The accretion disc plays an ***important role*** in the emission of jets. The ***key point*** seems to be the ***collapse*** of the disk. The theory of accretion discs predicts that the ***disc becomes unstable*** when it overheats and becomes optically too thick (Eddington limit) such that the thermal energy can no longer escape in the form of electromagnetic radiation [Mag01]. At this points the disk collapses and the inner part of the accretion disc falls into the black hole [Mei01].

Recently, hints have appeared of a ***connection*** between the ***collapse*** of the accretion disc and the ***ejection*** of a jet [Mar02]. This phenomenon has been observed in the case of 3C120, a radio quasar. The collapse of the accretion disc has been associated with the

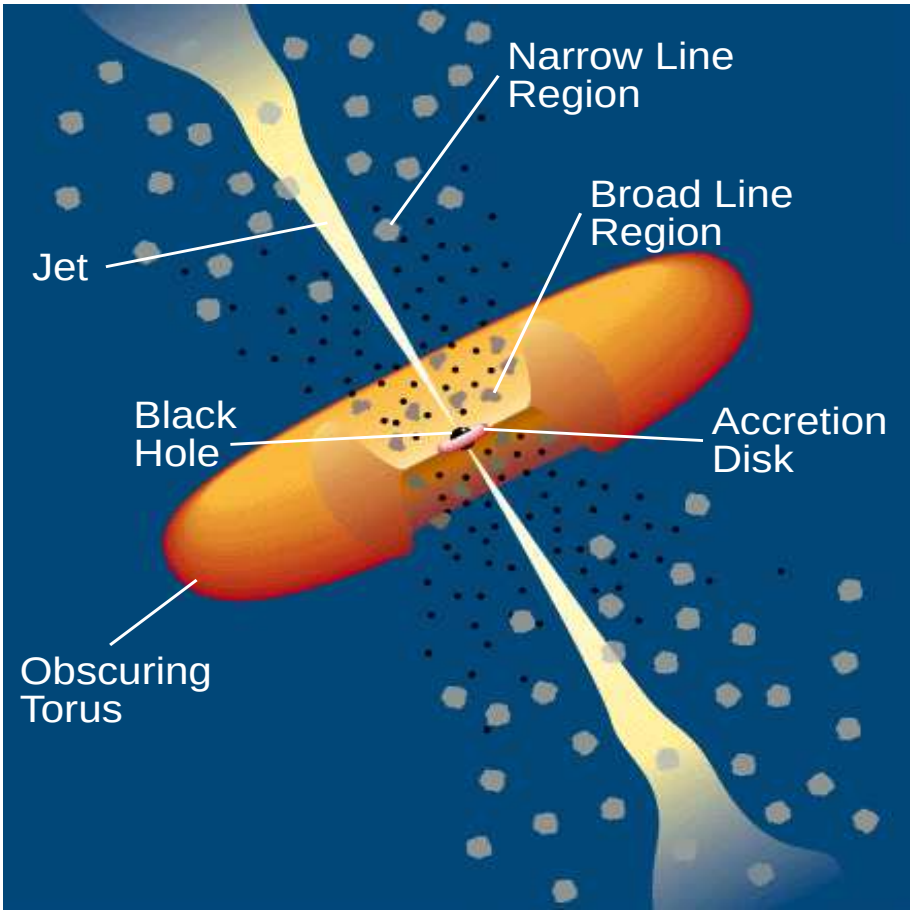


Figure 1.3: The AGN model is supposed to unify the different classes and subdivisions of AGN into one general picture. At the center of each AGN is a ***super-massive black*** hole between 10^6 to 10^{10} solar masses surrounded by a ***highly relativistic rotating accretion disc*** [Liv02]. Due to ***friction*** the disc heats up to temperatures in the order of keV, thus transforming the gravitational potential into thermal energy. The plasma emits a ***thermal spectrum*** which peaks in the (soft) ***x-ray*** region. Photons emerging from the disc and ***excite*** atoms and molecules in gas clouds, either close to the disc, which yields strong Doppler broadened optical emission lines, or further away from the torus, which produce narrow lines (because they are colder and have smaller velocities). The different spectral differences are believed to originate from ***different observation angles*** and also from different physical parameters, like the ***accretion rate*** and the ***spin*** of the black hole. Sometimes there are strongly collimated, highly relativistic jets or blobs emitted perpendicular to the accretion disc from the poles of the black hole. The ***radio emission is related to synchrotron emission*** of relativistic (i.e. electrons) particles in the jet (Picture taken from [Pad95]).

so-called ***x-ray dip,*** which is a spontaneous and significant decrease in the thermal (soft) x-ray emission of the disk. The soft x-rays have been clearly identified by an iron emission line, believed to originate from the inner part of the accretion disc. The dip is only seen in the soft x-ray emission. A ***clear correlation*** between the ***x-ray dip*** and the ***ejection*** of a radio blob has been observed.

To get a better understanding of the acceleration mechanism of jets, a ***laboratory experiment*** was setup [Hsu02]. The system of a central object and an accretion disk was simulated by a plasma (in form of a disc) together with a magnetic field generated by a magnet and an electric field between the central object (a round electrode) and the disc (an electrode ring). The jet was ejected when the electric field was switched on. In this experiment the jet was ***magnetically driven*** and also magnetically ***collimated.*** The experiment also showed that jets with helicity appear naturally in such a system (See Fig. 1.4).

Furthermore, ***general relativistic magnetohydrodynamic (MHD) 3-D simulations*** of the accretion-disk/black-hole system have shown that jets are a natural consequence of a rotating disk in the ***presence*** of a magnetic field [Mei01, Kud99, Koi02]. The simulations seem to suggest that mainly ***very fast rotating*** Kerr black holes (mass M) with spins from $a = 0.90\,M$ up to $a = 0.95\,M$ (maximally rotating is $a = 1.0\,M$) are able to accelerate jets up to Lorentz factors of $\delta = 10$. The accelerated plasma most likely originates from the ***outer halo*** of the accretion disk. The jet is ***magnetically driven*** and and also magnetically ***collimated***. The particles accelerating from the plasma are ***most likely electrons***. Heavier particles and ions are slower and remain in the disk halo [Mag01]. This would imply that the jet mainly consists of ***electrons*** and ***electromagnetic radiation.***

Evidence for the ***collimation*** of the jet at a distance to the nucleus between 30 and 100 Schwarzschild radii was recently observed in the radio galaxy M87 by J. Biretta et al. [Bir02]. During jet collimation, very close to the black hole, the jet opening angle appears to increase while at larger distances it decreases.

Other theories claim that the black hole acts like a ***huge dynamo*** in a magnet field generated by the accretion disc. The event horizon behaves like an electric conductor and builds up an electric field between the equator of the black hole and its pole. The static magnetic and electric field densities (***Poynting flux***) reach such a strength that e^+e^- pairs are created ***out of the vacuum***. These electrons are accelerated by the electric field and focused by the magnetic field [Mag01]. Such jets are purely electromagnetic.

Future measurements and simulations will clarify the question whether the jet only consists of electrons or if hadrons are present as well.

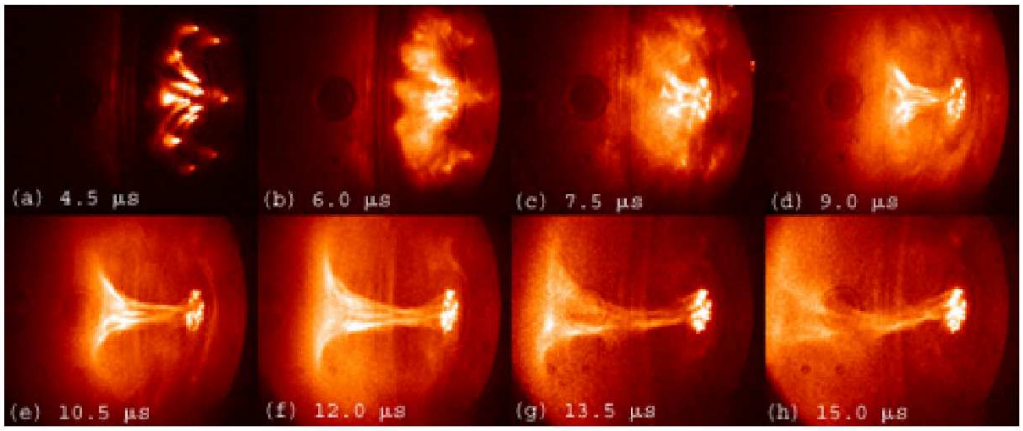
1.2.3 Jet models and the SSC model

It has been tried to explain the γ-emission by two different models:

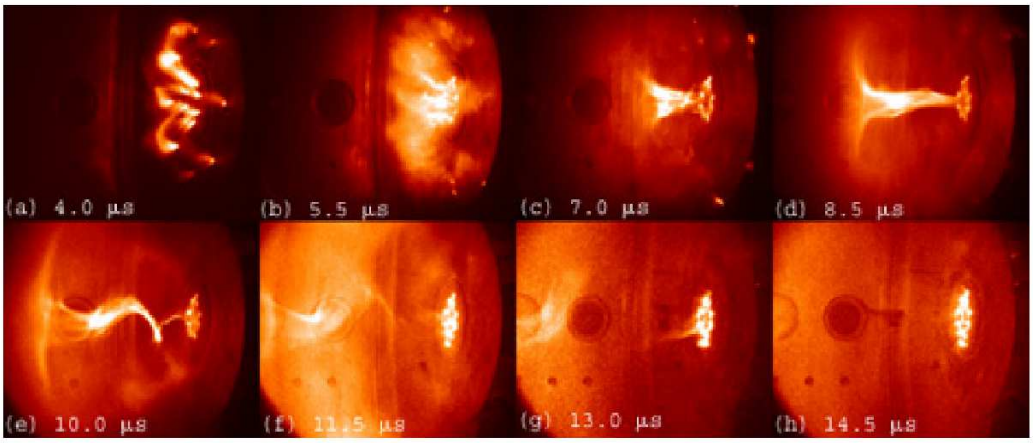
1. The ***synchrotron self Compton (SSC) model*** attempts to explain the production of very high energy gammas (TeV range). An elegant and apparently natural way to explain the existence of very high energy gammas is by inverse Compton upscattering of soft photons that are produced by the same high energy electrons that are involved in the scattering by synchrotron radiation.

2. In ***hadronic jet models*** the jet contains hadrons (e.g. protons) which produce π^0's in collisions. These π^0s decay into two γs and the high energy γ emission can be explained in this way. This model also would explain the ***existence*** of hadronic high energy cosmic rays.

Even that it seems very ***probable*** that a jet also contains a hadronic component, in the moment it seems that the (at least the pure) hadronic model has been ***discarded*** because the shape of the observed γ-spectrum to its theoretical prediction. On the contrary, the SSC model fits the shape of the observed spectrum very well. In this work I will restrict myself to the SSC model only.



(a) Development of a jet without helicity



(b) Development of a jet with helicity

Figure 1.4: Images of a laboratory experiment to designed to simulate jet emission which show the expulsion and development of a jet in a series of pictures. Plot a) shows a straight jet and plot b) shows a jet which develops a helical instability. Taken from [Hsu02].

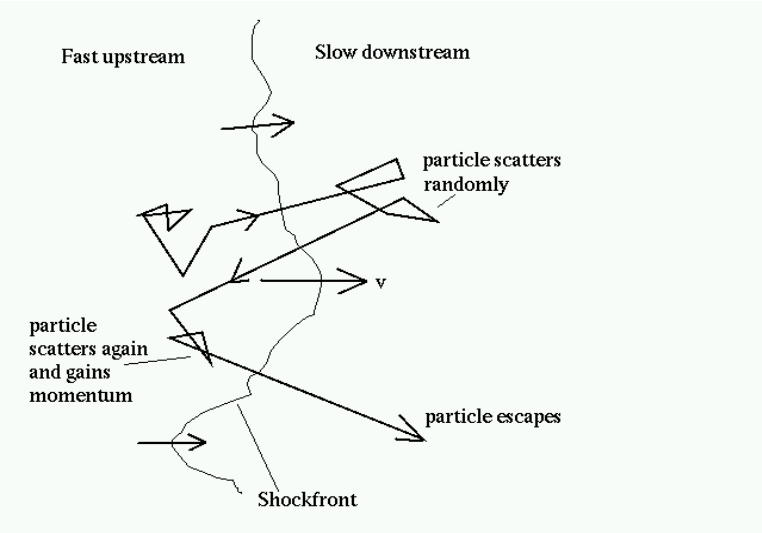


Figure 1.5: Illustration of the shock acceleration mechanism. The particle is scattered several times, forth and back over the shockfront. In each cycle it gains energy due to the movement of the shockfront.

The mechanism necessary to efficiently accelerate charged particles (e.g. electrons) to extremely high energies (> 1 TeV) is believed to be the so-called shock wave acceleration [Gal02].

Shock wave acceleration model

The just mentioned ***shock waves*** or shock fronts are naturally present in jets and super novae remnants (SNR). Particles are randomly ***scattered by local magnetic irregularities*** (Alfvén- and hydromagnetic waves [LON97/2]), thus passing many times from downstream of the shock front into the upstream interstellar medium (or if the shock front is inside the jet, from the downstream shock over to the slower upstream plasma in the jet) and back (see fig. 1.5). At ***every cycle*** the particle will ***gain energy*** due to the fast movement of the downstream shock [Gal02]. For nonrelativistic shocks the following formula shows the energy gain compared to the initial energy, averaged over all inclination angles.

$$\left\langle \frac{E_f}{E_i} \right\rangle = 1 + \frac{4}{3}\beta_{rel} \tag{1.3}$$

β_{rel} is the velocity of the jet in units of c. The particle has a ***probability to escape*** the acceleration front volume by being scattered away from the shock front. Taking this probability into the calculation the resulting spectrum has a ***power law*** shape. In nonrelativistic shocks the angular distribution of the movements is uniform and for a strong nonrelativistic shock this leads to power law spectra with spectral index $\alpha = 2$.

$$\frac{dN}{dE} \propto E^{-\alpha} \tag{1.4}$$

For ultra relativistic shocks, which is the case for jets emitted by AGNs, the escape probability and the average energy gain per cycle is more complicated due to the relativistic movement. The angular distribution of the high energy charged particles is not isotropic anymore. In [Gal02] it has been demonstrated, that by assuming different scenarios, the spectral index for ultra relativistic shocks is in the range of α=2.2 to α=2.3.

A very important question concerns the ***maximum energy*** that can be reached by this mechanism. The hard limit of the maximum energy depends mainly on two aspects.

1. The radius of the emission region must be ***larger*** than the Larmor radius of the particle.
$$E_{max} \simeq qB\gamma R_{jet} \tag{1.5}$$
B is the magnetic field in the jet, R_{jet} is the radius or jet scale, q is the charge of the accelerated particle and γ is the relativistic gamma factor of the particle. According to Equ. 1.5 for a jet radius of $R_{jet} = 0.01$ pc = $3 \cdot 10^{14}$ m, a gamma factor of $\gamma = 10$ and a magnetic field of B = 0.1 G the absolute maximum energy is in the range of 10^{14} eV or 100 TeV. These numbers give the order of magnitude that is generally expected in such jets.

2. The electrons are constantly being ***cooled down*** due to two effects and limit the maximum reachable energy if the cooling is becomes faster than the acceleration:
   (a) Emission of ***synchrotron radiation*** (which depends on the magnetic energy density u_B in the jet)
   (b) ***Inverse Compton*** scattering (which depends on the soft photon density u_{soft}).

The SSC model and relativity

As already mentioned, the idea behind the Synchrotron Self Compton model is that the observed high energy γ's are produced by an ***inverse Compton up-scattering*** of low energy photons (UV and soft x-rays) by high energy electrons. The required dense ***soft photon field*** is assumed to be produced via ***synchrotron emission*** by the same high energy electrons by magnetic irregularities in the jet (the relevant parameter is the magnetic energy density u_B). The SSC model is described in detail in Appendix A.

As an alternative (which is not SSC), the soft photons could also originate from the ***accretion disc*** if the latter is not too far away from the blob. This possibility will be discussed in the last chapter.

The ***typical SSC*** model spectrum consists of two emission peaks (see Fig. 1.6). One s***ynchrotron peak*** (the soft photon field) which peaks in the UV or soft x-rays region and one ***inverse Compton peak*** which peaks in the GeV-region.

Due to the ***relativistic*** motion of the source (along the jet axis), the emission intensity is ***boosted*** in the direction of motion and becomes stronger and more energetic for an observer for whom the jet approaching as compared to an observer in the moving (blob) frame [Mag01]. The so-called beaming effect is characterized by the Doppler factor δ:
$$\delta = \left[\gamma\left(1 - \beta cos\vartheta\right)\right]^{-1} \tag{1.6}$$
where β is the velocity in units of c with
$$\beta = \sqrt{1 - \frac{1}{\gamma^2}} \tag{1.7}$$
where γ is the relativistic Lorentz factor of the motion and the observer sees the jet at an angle ϑ.

The ***Lorentz-transformations*** of the local comoving jet system (whose observables are primed here, eg. dt', $I'(\nu')$) relative to the distant observer system (whose observables are unprimed, eg. dt, $I(\nu)$) yield:

- All longitudinal ***length scales*** and ***time scales*** must to be divided by the Doppler factor δ:
$$dt = \delta^{-1}\, dt' \tag{1.8}$$
$$dx = \delta^{-1} dx' \tag{1.9}$$

.

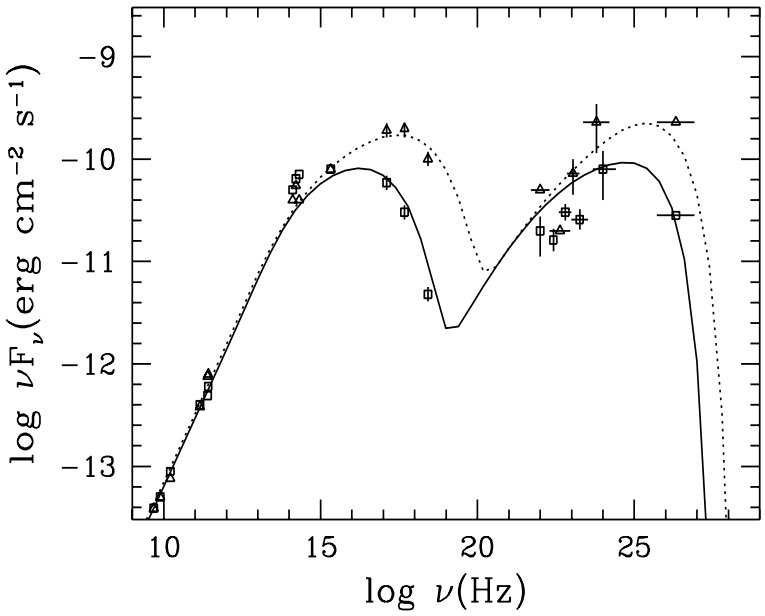


Figure 1.6: This figure shows an example of the double peak structure of a possible SSC model. Here the low and high states of the multi-wavelength spectrum of Mkn 421 have been fitted with a SSC model using a laminar geometry for the emission region. Taken from [Mas99]. Parameters used: variability time scale t=500s, with a Doppler factor of $\delta = 20$ (for low state, while for high states δ is by a factor four higher), with a magnetic field of B=0.4 G, and a maximum electron energy of $\gamma_{max} = 1.4\,10^5$).

- The ***intensity*** scales with δ^3 as:

$$I(\nu) = \delta^3 I'(\nu') \tag{1.10}$$

where $\nu = \delta \cdot \nu'$ and ν is the photon frequency. The opening angle of the light cone due to relativistic beaming is

$$\theta \simeq \tan\theta = \frac{1}{\gamma} \tag{1.11}$$

where theta is the opening angle of the cone and γ the Lorentz factor of the blob. Any (γ-) radiation emitted by the jet is beamed in forward direction.

1.2.4 Jets of Mkn 421 and Mkn 501 as seen by microwave telescopes

The jet properties are discussed on the basis of the two AGNs ***Mkn 421*** and ***Mkn 501***. The former is subject of this thesis and the latter will be used later for comparison.

The contemporary picture of a jet is a blob of relativistic particles that is ejected by the AGN. The ***Lorentz factor*** has been estimated for the case of many Blazars and radio galaxies by measuring the ***superluminal motion*** of the blob seen by radio telescopes. Radio telescopes have the necessary angular resolution to a precision of miliarcseconds, a resolution that is impossible to achieve for optical wavelengths. The general conclusion is that the Doppler factor of most blazars and radio galaxies is typically $\delta = 10$ [Pad2/99].

Superluminal motion is a ***natural effect*** of relativity in which the blob moves at angles of approximately 30°-60° in direction to the observer, who measures an apparent speed that is faster than light [Mag01]. The observed speed of the motion is:

$$\beta_{ob} = \left| \frac{\vec{n} \times \left(\vec{\beta} \times \vec{n} \right)}{1 - \vec{\beta} \cdot \vec{n}} \right| \tag{1.12}$$

$$= \frac{\beta \sin\phi}{1 - \beta \cos\phi} \tag{1.13}$$

where $\overrightarrow{n}$ is the observer direction, ϕ is the observation angle, $\overrightarrow{\beta}$ the velocity of the blob in the observer frame and β_{ob} is the apparent velocity observed by a telescope.

It is believed that the jet has a very high Lorentz factor ***immediately after*** its expulsion and collimation. Later it rapidly ***cools down*** by emission of synchrotron and inverse Compton radiation. The Lorentz factor of the blob decreases with increasing distance from the nucleus. At high Lorentz factors it radiates predominantly in the form of hard x-rays and (i.e. BLAZARS) in γ-rays. Eventually, when the jet is much further away from the nucleus and it has cooled down significantly, the jet emits synchrotron radiation in the microwave energy range at which time it transparent to microwaves . ***Only then*** will it become ***visible in the radio*** frequency range [Mar02]. At the point when the jet becomes visible to radio telescopes its distance from the AGN is already hundreds of parsecs.

Both objects, Mkn 421 and Mkn 501, have been observed with ***radio telescopes*** (see Fig. 1.7). As previously explained, it is assumed that the jets of these two objects move in the direction of the observer within an angle of $\theta \sim \frac{1}{\gamma} \sim 0.1\, rad \sim 6^\circ$ ($\gamma = 10$) from the fact that we see γ-radiation that is emitted in the forward direction. Unfortunately, the exact observation angle is ***unknown***. However, we do know that the movement is not exactly in the line of sight because moving blobs with superluminal velocity have been observed. These blobs are relatively slow. Mkn 421 blobs exhibit a velocity of only $v = 2 \pm 0.1\, c$ and Mkn 501 blobs have velocities of approximately $v = 2.5 \pm 0.1\, c$ [Mar99]. These ***slow superluminal velocities*** of the radio blobs have two explanations:

1. The jet initially had a ***higher*** bulk Lorentz factor of close to ten when it radiated x-rays and γ-rays (by synchrotron and γ-radiation emission) and subsequently ***cooled down rapidly***, more rapidly than other typical radio galaxies. Later the observed radio blob only shows a Lorentz factor ***of approximately 3***.

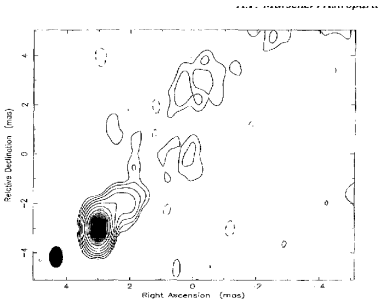
2. The ***observation angle is very small*** such that the superluminal motion has only small values between 2.0 and 2.5. This would allow a high Lorentz factor of the blob ***close to 10***.

It should be mentioned that the jet structure of both objects is significantly different. This can be seen for example in the magnetic field (can be seen in the polarization of the radio emission [Mar99]) which is perpendicular to the jet in case of Mkn 501 and for Mkn 421 the magnetic field is parallel to the jet.

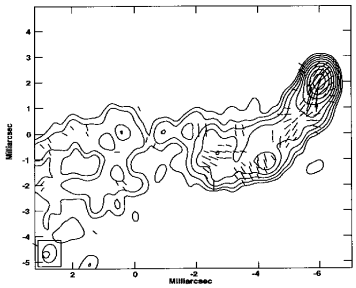
1.2.5 Fast variablity and other properties of AGN flares

As of today, several Blazars which emit gamma rays in the TeV range have been detected. The most significant of these are Mkn 501, Mkn 421, 1ES1426, 1ES2344 and 1ES1959, while astronomers have discovered about 60 AGN in the GeV range mainly by the EGRET satellite (see Fig. 1.8). All the known GeV Blazars are radio load. The radio emission is assumed to be synchrotron radiation of very high energy particles (mostly electrons but with a possible (weaker) contribution from protons) in the jet. The synchrotron emission extends up to the soft x-ray where it can be detected with satellites like ASCA or RXTE. Synchrotron radiation hints to the SSC model.

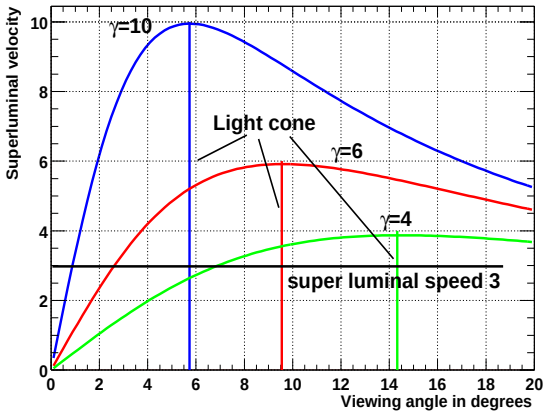
A very important feature of the AGN flares is its extremely fast variability. Mkn 501, assumed to be a black hole of 10^8 solar masses, and Mkn 421, smaller with approximately 10^6 solar masses are rather extended objects with Schwarzschild radii of the order of 1 AU and 1/100 AU respectively and with distances from the black hole to the accretion disc between 10 and 100 Schwarzschild radii. The distances from the black hole to the jet extend even further, with Schwarzschild radii at the order of 10^4 to 10^{10}. ***The observed variability is very rapid***. Typical timescales are ***a few hours for Mkn 501*** and about ***15***



(a) Mkn 421



(b) Mkn 501



(c) Superluminal speed for different Lorentzfactors

Figure 1.7: Shown here are VLBA radio images of the two objects Mkn 421 and Mkn 501, taken at 15 GHz. Picture a) shows Mkn 421, taken on July 1997 and picture b) displays Mkn 501, taken on August 1997. Mkn 421 is very core-dominated which could be due to a small observation angle. Mkn 501 shows a bending in the jet of almost 90°. Both radio jets show superluminal motion exceptionally slow for typical radio galaxies. Mkn 421 blobs have approximately a velocity of $\beta_{obs} = 2.0c$ and blobs of Mkn 501 have approximately $\beta_{obs} = 2.5c$. These plots were taken from [Mar99]. Plot c) shows the superluminal speed as a function of the observation angle (in degrees) for three different Lorentz factors (γ=4, γ=6 and γ=10). The vertical line shows the position of the light cone. Since from these two objects γ-emission is observed the actual observation angle must be smaller than this limit.

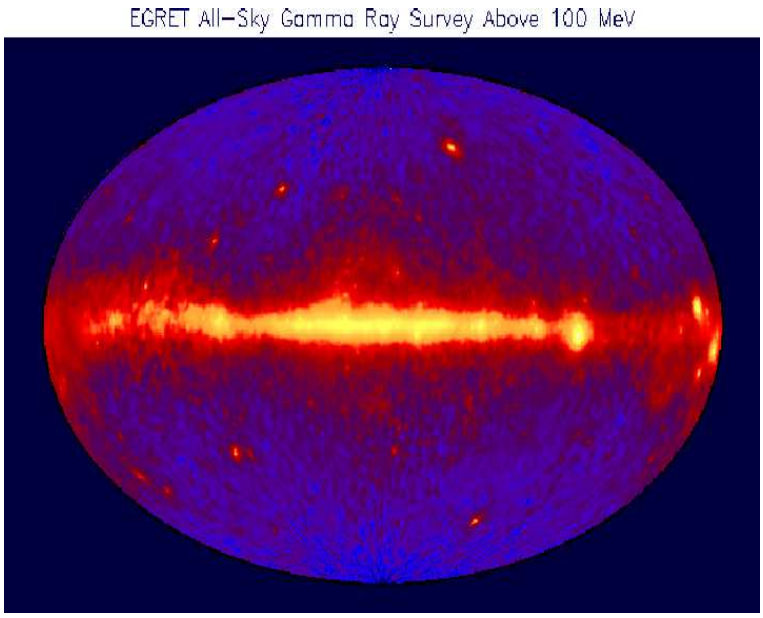


Figure 1.8: The as sky seen by the EGRET satellite above 100 MeV. Most of the emission comes from the region of the galactic plane but at higher galactic latitudes are some objects which are extragalactic. Many of them are Blazars.

minutes for Mkn 421. This can be partially explained by a high Doppler factor (Equ. 1.6) of the relativistically moving source and by small emission regions (which are tens of AU (1 AU$\simeq$ 15 light minutes)) .

But even with small emission regions and high Doppler factors ***common SSC models have problems*** to explain this fast variability of Mkn 421. New theories try to explain this behavior. They range from small ***conical sub shock fronts*** inside the jet (Model for Mkn 421 [Sal98]), ***laminar rather than spherical emission regions*** [Mas99], and ***modulation*** of the soft photon field for Compton up-scattering via a hot spot in the accretion disc [Bed96].

It should be possible to differentiate between these models by examining the ***correlation*** of x-ray flux and light curve measurements of x-ray satellites and the GeV-TeV emission. If the soft photons for up-scattering come from the accretion disc, there should only be a very weak correlation of the flares in the UV/x-ray and the TeV range. This would point to a modulation version.

If there is a ***strong correlation***, like in the case of Mkn 501, this points toward the ***SSC model***, possibly with ***conical shocks*** or ***thin laminal shocks*** traversing the jet (See Fig. 1.6).

A very precise measurement of a ***flare of Mkn 421*** in the x-ray region (see Fig. 1.9) has been recorded by ***ASCA in April 1998*** which seems to show a correlation between x-ray flux and γ-ray flux. The x-ray and TeV curves of Fig. 1.9 match perfectly, when superposed. In the analysis chapter the correlation between the x-ray flux as measured by ASCA in 1998 and various TeV instruments will be examined in detail. The ASCA x-ray flux measurement seems to show two components. A slow flare component ($\sim$7 day duration) and a fast (sub-day) flare component.

1.3 The Gamma Ray Horizon

The universe is filled with isotropically distributed diffuse photons of low energy from the microwave energy range (i.e. cosmic microwave background or CMB) peaking at 2.7 K up

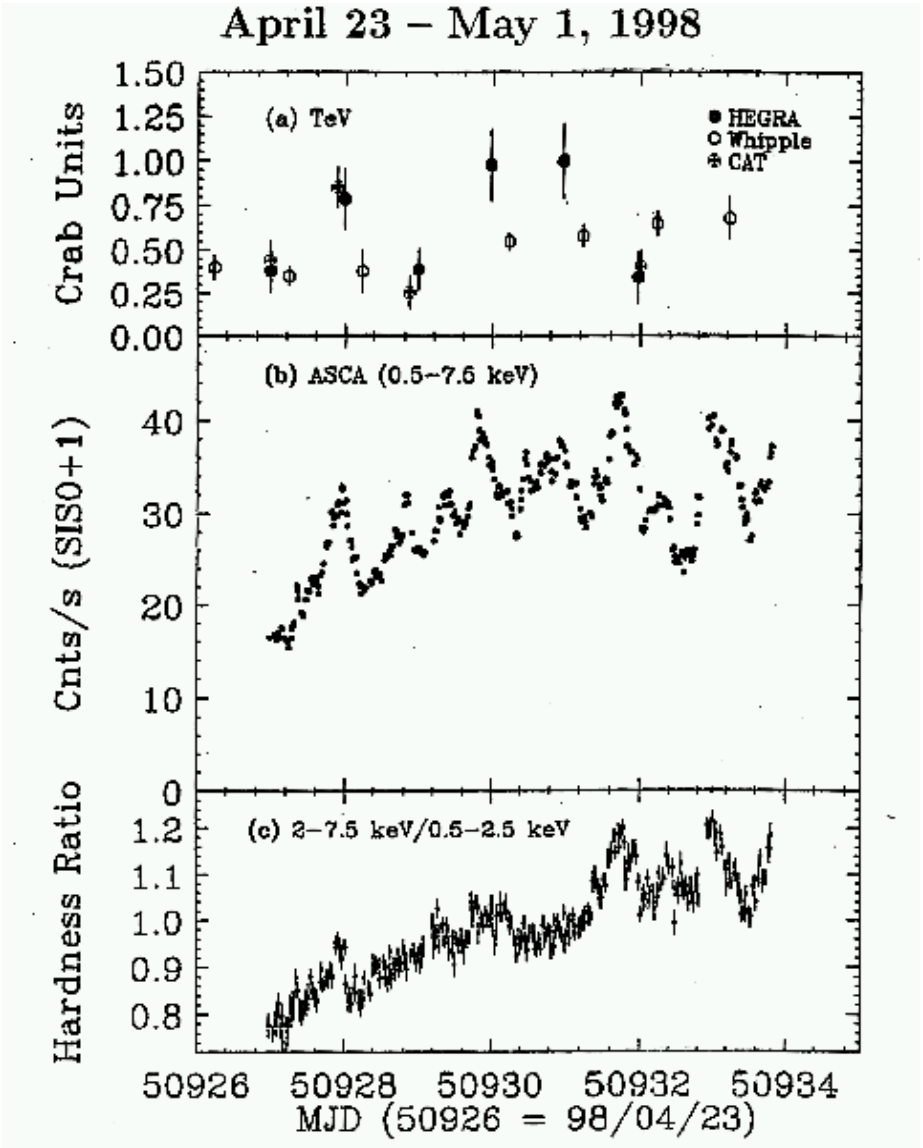


Figure 1.9: Multi-wave length observation of Mkn 421 in 1998 with ASCA in the x-ray region and with several Cherenkov telescopes in the TeV energy range. The integrated TeV flux is measured in units of the (constant) crab flux and the x-ray flux (rate) in counts/sec. When the x-ray curve is superposed on the TeV curve it can be seen that the TeV curve follows the x-ray curve. This will be discussed in detail in chapter B.

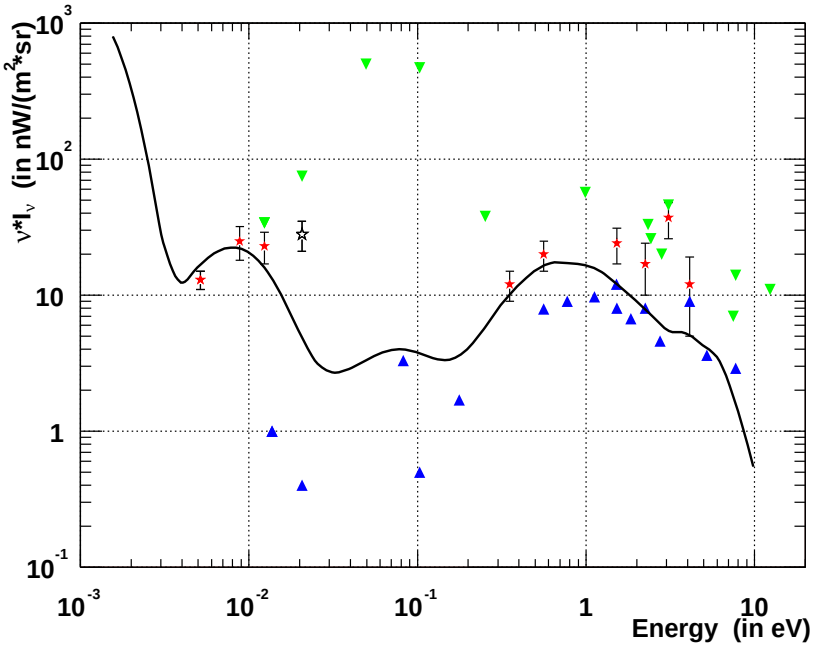


Figure 1.10: Actual measurements of the cosmic infrared background. The green points are upper limits, the blue ones are lower limits from galaxy counts, the red ones are trusted measurements and the hollow black one is a tentative measurement. The data has been taken from [Hau01]. The model is from T. Kneiske and K. Mannheim [Kne02]. The increase of the energy density on the left side is the CMB.

to the infrared (i.e. ***cosmic infrared background*** or CIB) and the visible light range. The interaction of the very high energy (VHE) γ's with the low energy background radiation (VLE photons) plays an important role in gamma ray astronomy.

This interaction ***limits*** the observable energy range as a function of γ-energy. Therefore one would expect ***cutoffs*** in the spectra. This attenuation effect of the VHE γ on their way to the earth, which depends on the red shift of the source, is called ***gamma ray horizon*** (a more precise definition will be given later, Equ. 1.23). The source of the interaction is electron positron pair production from high energy and low energy photons.

$$\gamma_{VHE} + \gamma_{VLE} \rightarrow e^+ + e^- \tag{1.14}$$

The VLE photons have different origin. The 2.7 K microwave background radiation is a remnant of the Big Bang with a thermal spectrum from 10^5 μm to 10^3 μm. Another important contribution from the ***far infrared (***100 μ***m) to the UV (0.1*** μ***m)*** comes from redshifted star light that has been occurring throughout the history of the universe. Star formation processes in the early universe play an important role here. Thus the precise measurement of the spectrum of these background photons provides detailed information about the history of our universe.

Unfortunately the direct measurements of the extragalactic infrared background by satellite borne instruments are ***very difficult*** because of foreground infrared light in our galaxy, i.e. dust reflecting the (infrared) light from the sun and stars inside the galaxy. The actual measurements of the infrared background can be seen in Fig. 1.10. The numbers have been taken from [Hau01]. The region between a few μm and 100 μm has not been measured at all. Only upper and lower limits exist. The lower limits come from galaxy counts and are hard limits [Hau01]. The plotted curve shows a models developed by T.

Kneiske and K. Mannheim [Kne02]. This model is quite complex and involve convolutions of measured star formation rates, initial mass functions and the history of dust and light densities.

In order to calculate the absorption that VHE photons experience we need to ***fold*** the cross section of the pair-production process with the low energy photon density. ***The cross section*** for the process described in Equ. 1.14 is (in cm^2) [Ste95]:

$$\sigma\left(E\left(z\right),\varepsilon\left(z\right),x\right) = 1.25\cdot 10^{-25}\left(1-\beta\right)\times \tag{1.15}$$

$$\left[2\beta\left(\beta^2-2\right)+\left(3-\beta^4\right)\cdot ln\left(\frac{1+\beta}{1-\beta}\right)\right] \tag{1.16}$$

where

$$\beta=\sqrt{1-\frac{2\left(m\,c^2\right)}{E\varepsilon x(1+z)^2}} \tag{1.17}$$

E is the energy of the gamma ray photon, ε is the energy of the soft photon, z is the redshift of the source and $x=(1-\cos(\vartheta))$ is the angle between the photon directions. ***The threshold energy condition*** is

$$E\varepsilon x\left(1+z\right)^2>2\left(mc^2\right)^2 \tag{1.18}$$

The cross section, plotted in the range between 100 μm and 0.1 μm, for a 1 TeV, a 3 TeV, a 10 TeV and a 30 TeV photon from a head on collision (solid line) and averaged over all angles (dashed line), together with the measurements of the infrared background can be seen in Fig. 1.11. This time the photon density instead of the energy density has been plotted, because this is the quantity which determines the absorption. This plot illustrates which soft photon range interacts with which high energy photon. The higher energy photons (10 TeV and 30 TeV) fall into an energy range of the cosmic ray background which has not been measured yet. The measured spectrum of Mkn 421 puts strong constraints and upper limits on the soft photon density in that range.

To get the optical depth τ (attenuation to $e^{-\tau}$ of the original flux) for one specific energy, one must ***fold*** the cross section with the cosmic infrared spectrum and average over the whole solid angle Ω of photon collision angles and integrate this result over the redshift z. Since the object of this work (Mkn 421 with z=0.03, z$\ll 1$) is relatively close to the earth (440 million light years away), it is not necessary to integrate over z, which simplifies the relation.

The absorption probability per unit length (in cm) is [Ste01, Bla01]:

$$\frac{d\tau}{dl}=\frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin\vartheta d\vartheta\int_{\frac{2m^2c^4}{E(1-\cos\vartheta)}}^{\infty}d\varepsilon\frac{dn}{d\varepsilon}\sigma\left(E,\varepsilon,(1-\cos\vartheta)\right) \tag{1.19}$$

$\frac{dn}{d\varepsilon}$ is the ***differential soft photon density*** (in units of number of photons per $cm^3$and eV) and must be calculated from the cosmic infrared background spectrum (Intensity $\nu\frac{dI}{d\nu}=\nu I_\nu$ (in units of nW/(m^2sr))) in the following way:

$$\varepsilon^2\frac{dn}{d\varepsilon}=\frac{4\pi}{ec}\quad\nu I_\nu \tag{1.20}$$

whereas ε denotes the photon energy in eV. The distance to the object is, to first order, (for z$\ll 1$):

$$l=\frac{c\cdot z}{H_0} \tag{1.21}$$

$H_0=68\pm6\frac{km}{sec\cdot Mpc}$ is the Hubble constant (1 pc = $3.086\cdot10^{18}$ cm). Finally the attenuation of the flux becomes:

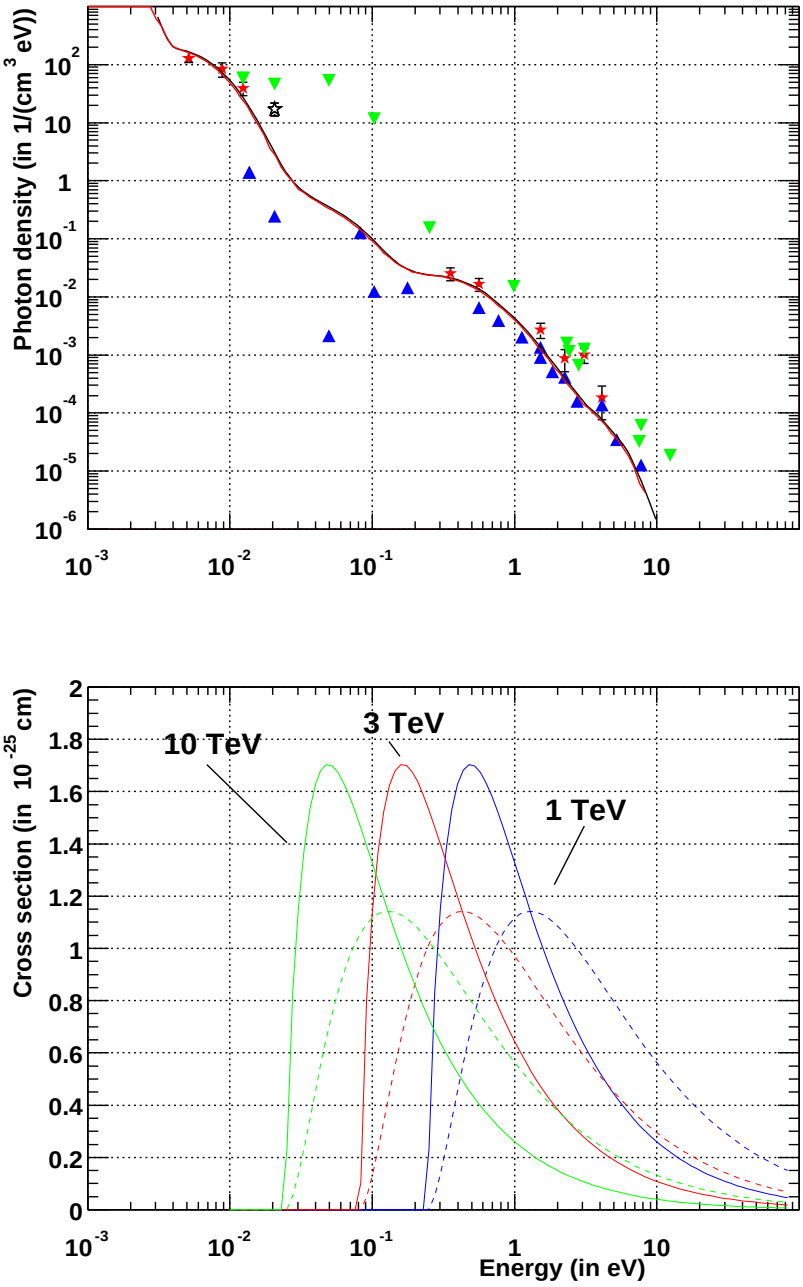


Figure 1.11: a) The upper plot displays, for comparison, ***the cosmic infrared background measurements*** in terms of differential photon density together with different models of [Kne01] of the CIB. The blue points are lower limits, the green ones are upper limits and the red points are accepted measurements. b) The lower plot shows the ***cross section*** for a 1 TeV (blue), a 3 TeV (red), a 10 TeV (green) and a 30 TeV photon from a head-on collision (solid line) and averaged over all collision angles (dashed line).

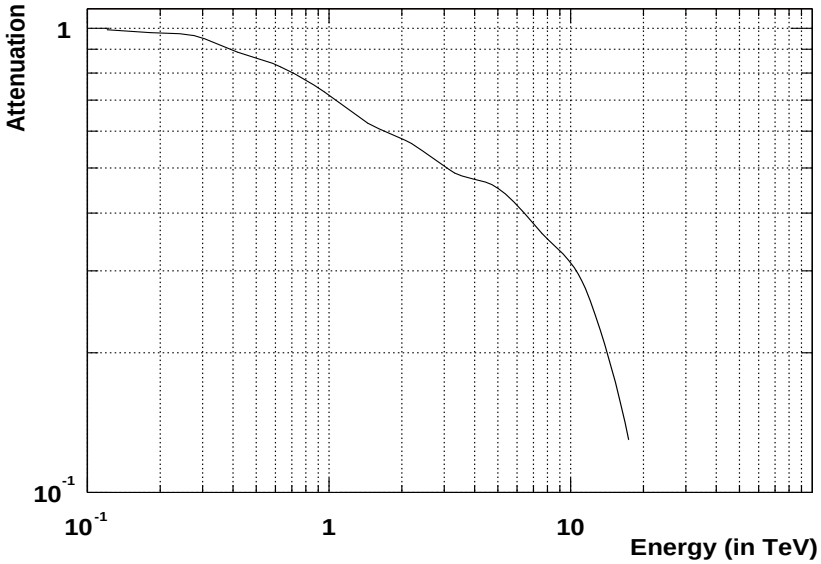


Figure 1.12: The ***attenuation*** of TeV gammas originating from Mkn 421 and according to the models [Kne02] shown in Fig. 1.11. According to these models the attenuation starts above 1 TeV but the effect becomes ***significantly*** strong after 10 TeV.

$$A = e^{-\frac{d\tau}{dl} * l} \tag{1.22}$$

Fig. 1.12 shows the ***attenuation*** of gammas originating from Mkn 421 at z=0.031 for the models shown in the Fig. 1.11. It should be noted that the ***cutoff*** that has been observed, for example, in the spectrum of Mkn 501 (z=0.034) [Ste01] can have two possible reasons. One is the ***absorption*** due to the cosmic infrared background. The other is that the cutoff could be an ***intrinsic feature*** of the BL LAC emission spectrum. As of today, it is not clear which of these hypotheses is actually responsible for the cutoff. It could even be a ***mixture of both***. These possibilities will be discussed later in the last Chapter concerning the analysis of the Mkn 421 flares of 2001.

The gamma ray horizon (Fig. 1.13) is defined as the red shift for which

$$\tau\left(E, z\right) = 1 \tag{1.23}$$

This is the distance z for which the attenuation becomes $1/e$.

1.4 Theory of air showers

There are two ways of detecting gamma photons. One is such that the gamma photons are detected directly in space. This is done via the x-ray and gamma-ray satellites. Satellites have a limited detection area of a few m^2 and have therefore a limited sensitivity. The second possibility uses the ***atmosphere*** as a part of the ***detector***.

The VHE gamma flux is so low, that huge detection areas ($\sim 10^5\ m^2$) are needed to collect a reasonable number of events per observation time. Above 10 GeV the collection area of current satellite borne γ-detectors is too small and only the method of ***ground based instruments*** can be used. This section is therefore dedicated to explain how the atmosphere reacts with cosmic rays.

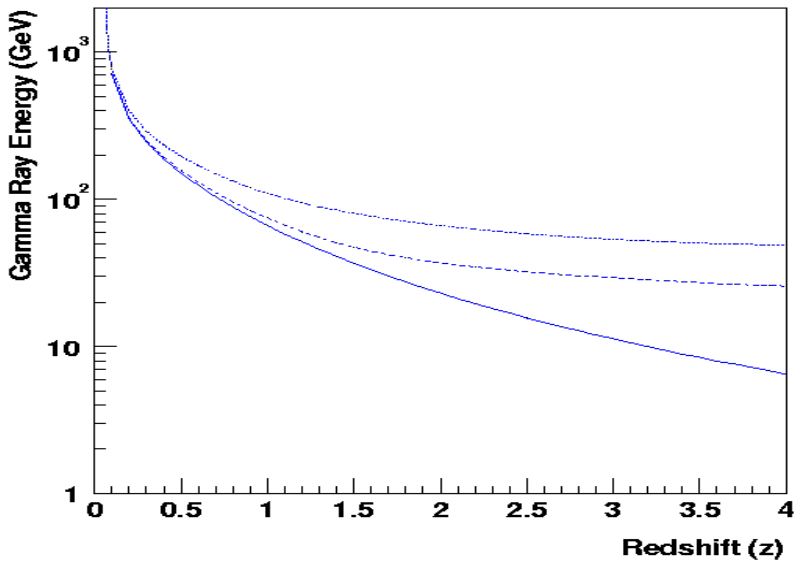


Figure 1.13: The gamma ray horizon plotted against the red shift. Taken from [Bla01]. As of today, it is not exactly known where the horizon is located. The different curves show the attenuation for different cosmic parameters and CIB models. The location of the gamma ray horizon gives important information about CIB and is a very actual field of research.

When cosmic rays enter the atmosphere, they generate so-called air showers. The cosmic ray particles interact with the molecules of the atmosphere by hadronic and electromagnetic interaction. ***Electrons, muons and γ's interact electromagnetically***, which means they generate secondary particles by pair production and bremsstrahlung. The secondary particles are mostly again non-hadronic particles as ***Electrons, muons and γ's***.

Hadronic CR, namely ***Protons and ionized nuclei*** interact via the ***hadronic interaction***, which means that they produce as secondary particles mainly π's, μ and K's. These latter particles either decay or produce more secondary particles by hadronic interaction. π^0 decay almost instantly into two photons and therefore ***feed the electromagnetic component*** of an hadronic shower.

In this way, a cascade of secondary particles is initiated. These highly relativistic particles, mainly electrons, emit Cherenkov radiation during their travel through the atmosphere. The air showers extend, depending on their energy, from the upper levels of the atmosphere down to sea level (see Fig. 1.14).

There are two types of air showers: The ***electromagnetic type***, which have ***no*** hadronic particles and the ***hadronic showers***, which contain all types of particles.

1.4.1 Electromagnetic cascades

The electromagnetic shower is easier to discuss mathematically because to first order it ***only*** contains ***electromagnetic particles***, consisting of ***electrons***, ***muons*** and ***photons***. Photons are produced by ***bremsstrahlung*** and by the ***annihilation*** of positrons. Electrons and positrons are produced by ***pair production***. The energy of the original photon is transfered to the secondary particles whose number ***increases*** continuously.

The particles ***lose energy*** by ***multiple scattering*** and also by ***ionization*** of the surrounding air molecules. The number of particles increases until the shower maximum. From that point on the average particle energy decreases because of energy losses due to ionization and bremsstrahlung. The cascade equations can be found in [Gai90]. For the

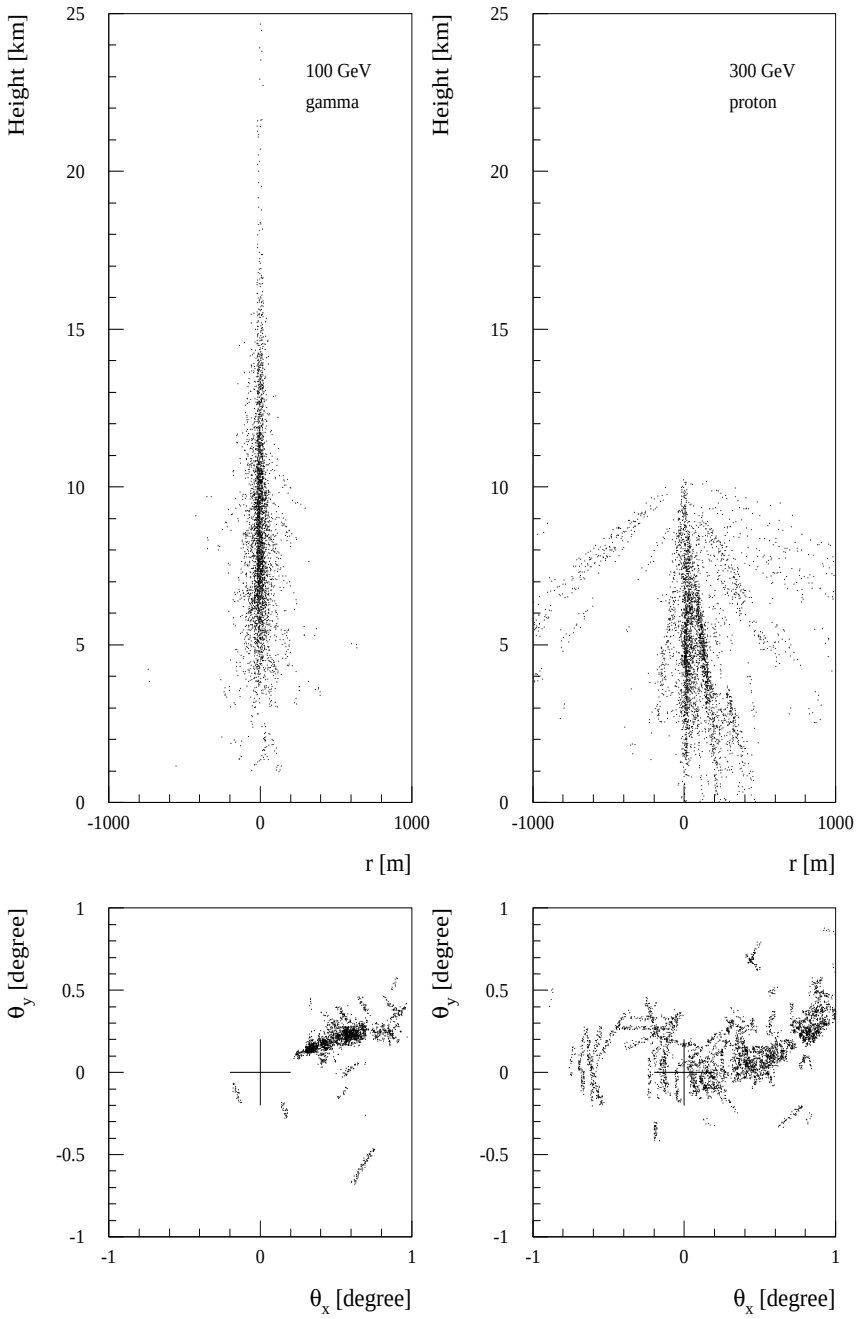


Figure 1.14: a) Top row: Images of two typical air showers, for a gamma photon (left side) and a proton (right side) , calculated by Monte Carlo simulation and 2) Bottom row: their Cherenkov image seen by a telescope, taken from [Kra01]

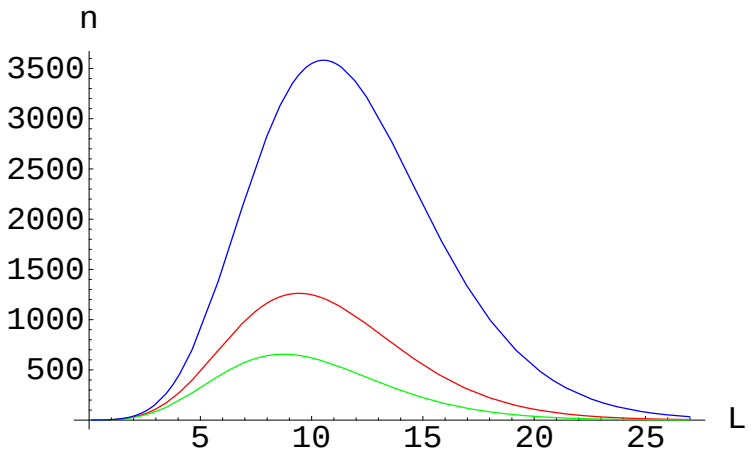


Figure 1.15: ***Electron number*** of a 0.5 TeV (green), of a 1 TeV (red) and a 3 TeV (blue) shower vs ***depth*** in units of radiation lengths.

case of electromagnetic showers solutions ***exist*** for the cascade equations which assume several approximations.

The ***mean free path length*** of an electron (distance at which the 1/e part of the electrons did not interact) is called the ***radiation length*** λ and has the size of $\lambda = 36.2 \frac{g}{cm^2}$. This quantity has to be divided by the density of the air to get the unit of a length (at sea level ρ = 0.0011 $\frac{g}{cm^3}$ it equates to approximately 350 m). For the following discussion four parameters are introduced:

- The ***depth T*** along the shower axis in radiation lengths λ..
- The ***shower age*** $s = 3/(1 + 2 * (y/T))$, which ranges from 0 to 2. The shower maximum is found at s=1.
- The ***critical energy*** E_c = 80 MeV, where the energy loss through ionization, equals the energy loss through particle multiplication.
- And a ***variable*** $y = \ln(E/E_c)$, where E is the energy of the initial gamma photon.

The approximate longitudinal electron number then becomes [Gai90]:

$$N_e(s) = \frac{0.31}{\sqrt{y}} e^{T(1-1.5*\ln s)} \tag{1.24}$$

The ***electron number*** vs ***depth*** for a 3 TeV, a 1 TeV and a 500 GeV shower is plotted in Fig. 1.15.

By introducing a very simple exponential atmospheric model (neglecting temperature changes), one obtains an approximate shower shape as a function of height. The following equation relates the depth T (in radiation lengths) with the height H (in m):

$$T = \frac{X_0}{\lambda \cos \vartheta} e^{-\frac{H}{H_0}} \tag{1.25}$$

Where $X_0 = 1013 \frac{g}{cm^2}$ is the column height of air at ground, $H_0 = 8400$ m is the height at which the atmospheric pressure reduced to 1/e from the one at ground and $\vartheta = 0$ is the inclination angle of the shower axis.

The resulting plot in Fig. 1.16 shows a 3 TeV, a 1 TeV and a 500 GeV shower with respect to height according to the simple model described above. The shower maximum for a 1 TeV shower can be seen at approximately 9500 m. The shower maximum for showers with higher energy ***reaches deeper*** into the atmosphere.

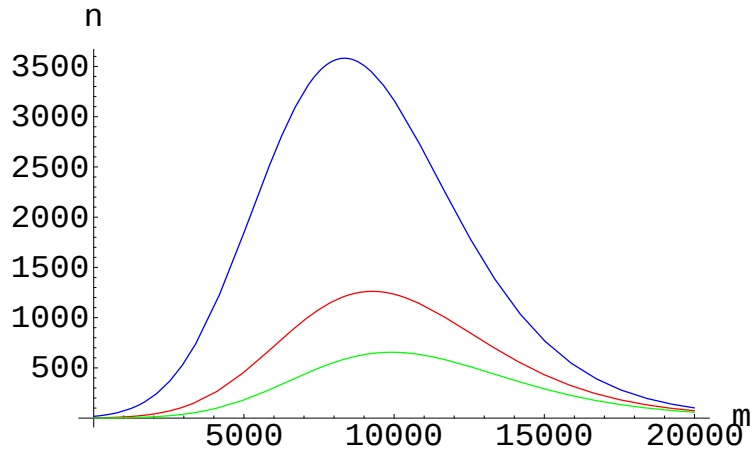


Figure 1.16: ***Electron number*** of a 0.5 TeV (green), of a 1 TeV (red) and a 3 TeV (blue) shower shower vs the ***height*** in m counted from ground. For higher energies the shower maximum ***reaches deeper*** into the atmosphere.

In order to understand the ***asymmetric shape*** of shower images seen by the camera, we analyze the ***electron density*** seen under an observation angle by the telescope (the Cherenkov image itself is far more complicated and will not be treated here). The transformation from the height coordinate to the viewing angle coordinate involves a ***tangent*** of that angle. This results in a ***very asymmetric*** shower image. Fig. 1.17 shows the same three showers developing at an impact parameter of ***90 m***. Here it can be seen clearly that the shower image in the telescope camera is very asymmetric. The shower maximum moves to larger ***distances*** for in the camera for higher energies.

The camera of the CT1 telescope has an opening angle of approximately ***1.4 degrees***. Comparing this with our image, it can be concluded that the recorded camera image will be ***truncated***. For these reasons, two new image parameters have been introduced, as will be explained in detail in a later chapter. They are: an ***asymmetry parameter,*** and a ***leakage parameter***, which describes the ***degree of truncation*** of the shower image due to a too small camera. The latter parameter will be used for energy estimation purposes.

There also exists an analytical solution for the ***lateral width*** of the shower. Unfortunately it is only valid for $1 \leq s \leq 1.4$ and is therefore ***not very useful*** for developing an analytical model for the shape of γ-showers. (Such a model might be useful for γ/hadron separation using log-likelihood techniques). It is the so-called ***NKG-formula*** displayed in Fig. 1.18:

$$xf(x) = K\, x^{s-1}\, (1+x)^{s-4.5} \tag{1.26}$$

where K is a normalization constant and

$$x = r/r_M \tag{1.27}$$

is the radial distance from the shower core in units of ***Moliere radii***

$$r_M \approx 9.3 \frac{g}{cm^2}$$

The following expression defines the normalization constant K in Equ. 1.26:

$$2\pi \int_0^\infty xf\,(x)\, dx = 1 \tag{1.28}$$

Finally, the ***lateral electron density*** $\rho_N\,(r,T)$ becomes

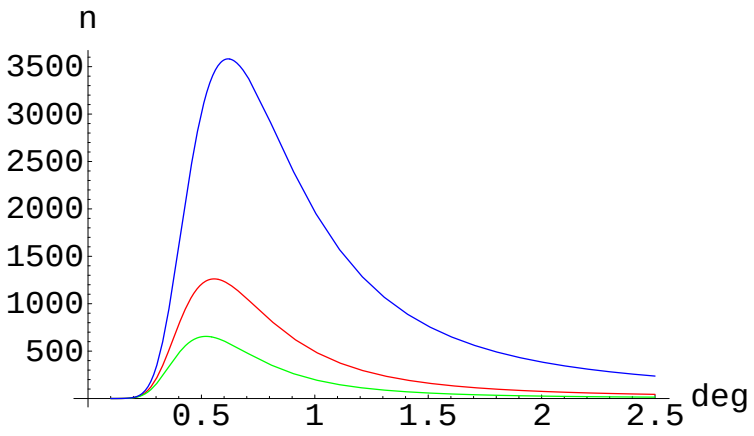


Figure 1.17: ***Electron number*** of a 0.5 TeV (green), of a 1 TeV (red) and a 3 TeV (blue) shower vs the ***viewing angle*** from ground in degrees. For higher energies the shower maximum is found at ***larger distances*** in the camera.

$$\rho_N\left(r,T\right)=\frac{N_e\left(T\right)}{r_M^2}f\left(x\right) \tag{1.29}$$

As a result, it can be stated that the ***lateral spread increases*** with the ***shower age***. From the view of the camera the image develops from the center (if the center of the camera points to the source) to ***bigger*** viewing angles while the lateral spread continuously ***increases***.

Hadronic showers are more complicated than electromagnetic ones and the cascade equations cannot be solved analytically. Usually ***Monte Carlo simulations*** are needed to understand the difference between electromagnetic cascades and hadronic cascades.

1.4.2 Hadronic cascades

A hadronic shower has ***three*** components, ***a hadronic***, ***an electromagnetic*** and ***a muonic*** one. The shower develops as follows: A high energy hadron interacts with the nucleus of an atom in the air and produces mainly pions (and a few kaons).

$$h+Nucleus\ \rightarrow\ m\pi^{\pm}+n\pi^{0} \tag{1.30}$$

The shower consists of a hadronic high energy core that continuously feeds the electromagnetic part because the π^0 instantly decay into two photons.

$$\pi^{0}\ \rightarrow\ \gamma+\gamma \tag{1.31}$$

Each high energy photon emerging from the ***hadronic core*** creates an ***electromagnetic sub-shower***. Lower energy charged K's and π's feed the ***muonic component*** by decaying into muons and neutrinos. At each hadronic interaction, ***approximately one third*** of the energy goes into the ***electromagnetic component***. Since the hadrons usually re-interact, a large part of the initial energy finally ends up in the electromagnetic part and is dissipated through ionization losses and Cherenkov radiation. The most numerous particles in a hadronic shower are therefore positrons and electrons. It should be noted that in a hadronic shower a ***sizeable fraction*** of the energy is transported away by invisible neutrinos and muons and so that ***only a fraction*** of the initial energy is ***deposited*** in the atmosphere. Therefore hadronic showers produce ***less*** Cherenkov light than their electromagnetic counterparts.

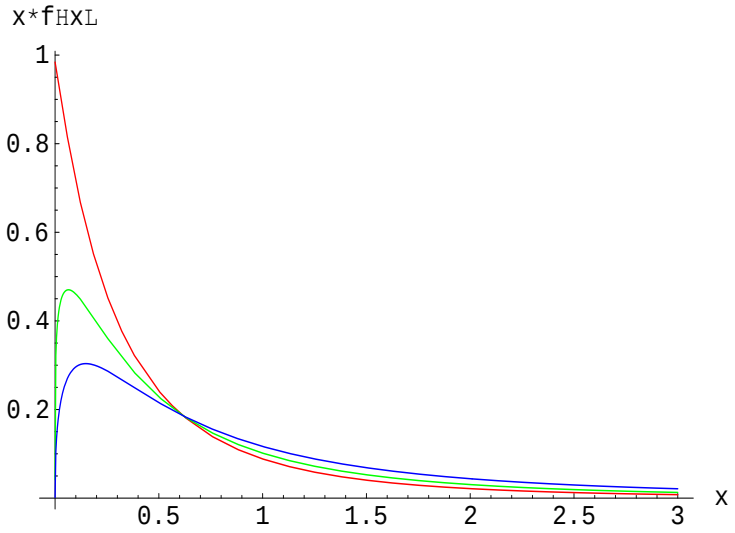


Figure 1.18: The function $xf(x)$ with correct normalization demonstrates the lateral electron distribution for s=1.0 (red), s=1.2 (green), s=1.4 (blue). The lateral distribution gets ***wider*** with shower age due to ***multiple scattering***.

A Cherenkov telescope sees the Cherenkov radiation basically only from electrons and positrons because a) they are more likely to be ***above threshold*** of Cherenkov light production and b) they form the ***overwhelming*** fraction of particles in a shower. The image of an hadronic shower seen by such a telescope is wider and larger and has much ***larger fluctuations*** than that of a pure electromagnetic shower (see Fig. 1.14). The fluctuations are larger because the hadronic radiation length is almost double the size of the electromagnetic radiation length: $\lambda_h = 70\frac{g}{cm^2}$. The ***lateral spread*** of the hadronic shower is mainly caused by the ***transverse momentum*** of the secondary hadrons after a hadronic interaction while in the case of electromagnetic showers the lateral spread is a function of ***multiple scattering*** resulting in a much slimmer shower.

Since one is interested only in gamma ray induced showers, the hadronic showers must be ***separated*** from the electromagnetic ones. This difference in the ***geometrical*** image structure, and to a certain extent the ***time*** structure, can be used to distinguish between the gamma induced events and the hadronic background, provided the instrument has sufficient resolution (e.g. a finely pixelized camera and sub-nanosecond time resolution).

It should be noted that an incoming high energy γ-photon (or a γ in the shower) can develop secondary hadrons with low probability (1%) via photoproduction which initiate a hadronic shower. Such showers, induced by a γ-photon, cannot be distinguished from a normal hadronic shower and therefore add to the background. However, most will be eliminated by selection cuts.

Below 100 GeV there exists a small but ***non-negligible contribution of electrons*** in the primary CR flux. These electrons initiate electromagnetic showers as primary γ's and are therefore indistinguishable from γ induced showers and so form a non-reducible background. However, like all the other charged CR, they are ***isotropically distributed***. It is therefore essential to achieve a good angular resolution to reduce the e^- background in the case of point source searches.

All the structure of extended air showers can be seen through the Cherenkov light emitted by the particles (with $\beta \geq \beta_c$) in the shower. Cherenkov telescopes are especially sensitive to the ***directionality*** of the Cherenkov light. This is an important difference to air ***fluorescence*** experiments.

1.4.3 Cherenkov light production

A highly relativistic particle emits Cherenkov light during its travel through the atmosphere [Jel58, Lon1/92]. Cherenkov light is always produced when the particle velocity is ***faster then the velocity of light*** in the air ($\beta \geq \beta_c$). The condition for Cherenkov light production therefore is:

$$\frac{c}{n\left(\rho,\omega\right)} < v = \beta c = (1 - \frac{1}{\gamma^2})c \tag{1.32}$$

where $n\left(\rho,\omega\right)$ is the refraction index of the atmosphere which ***depends*** on the ***density of the air*** and the ***wavelength***, β is the velocity of the particle in units of c and γ is the relativistic Lorentz-factor. The condition above also ***limits the spectral range*** of the Cherenkov-light since the refraction index of the air depends on the wavelength. The Cherenkov light is radiated in a light cone of opening angle

$$cos\left(\theta\right) = \frac{1}{\beta n\left(\rho,\omega\right)} \tag{1.33}$$

The maximum opening angle ***increases*** as the particle enters deeper in the atmosphere due to the increasing air density. This is illustrated in Fig. 19.The Energy radiated per unit length and unit frequency is [Lon1/92]:

$$\frac{dE}{dl} = \frac{e^2\omega}{4\pi\varepsilon_0 c^3}\left(1 - \frac{c^2}{n^2\left(\rho,\omega\right)\nu^2}\right) \tag{1.34}$$

The Cherenkov light seen from the ground is the ***superposition*** of all the light emitted in cones by all the electrons and positrons integrated over the whole shower development.

In order to determine the total light distribution on the ground, the ***lateral spread*** of the electrons and more importantly, the ***angular distribution*** $q\left(\beta,h\right)$ of the ***propagation direction*** of the electrons, which defines the ***direction of emission*** of the Cherenkov cone, has to be taken into account. Fig. 19 shows the distribution of Cherenkov light emitted by a 1 TeV shower.

.1 The principle of Cherenkov imaging telescopes

Air showers develop practically with the speed of light, resulting in very short Cherenkov light flashes. Typical numbers for the arrival time of Cherenkov photons are 2-4 ns for an electromagnetic shower. Hadronic showers have a wider time spread (10 ns - 15 ns) due to the development of many sub showers. Fig. 20 shows the Cherenkov photon density on ground for different CR particles. The ***photon density*** for electromagnetic showers on the ground ***scales to first order with the energy*** (for E > few GeV; the light output is used as the main energy estimator) while for protons this relation does not hold below 1 TeV. Reasonable mirror areas are needed to collect a sufficient amount of light for detection.

For example, for a ***1 TeV*** electromagnetic shower ***only about*** 100 photons/m^2 arrive the ground in the main impact parameter region of a radius of about 120 m, within the wavelength region of 300 nm (the ozone cutoff) to 600 nm (the sensitivity limit of the PMT). Assuming a collection mirror ***area of*** $\sim 10\,m^2$ for the CT1 telescope and taking into account losses in the optics, ***only about 800 photons*** will arrive to the (whole) camera during a few nanoseconds. During this time also light from the ***night sky*** (NSB) is continuously being collected by the mirror. On the Roque de los Muchachos in La Palma at 2300 m (the location of the CT1 telescope, the NSB accounts for approximately $4.3\ 10^{12}\ \frac{photons}{s\,m^2\,sr}$ [Sch00]. For $\sim 10\,m^2$ and a pixel diameter of 0.25° (corresponding to the CT1 telescope) this results in approximately ***0.6 photons per ns and per pixel.*** This number depends strongly on the sky area and weather conditions and can easily change by a factor of 2 or

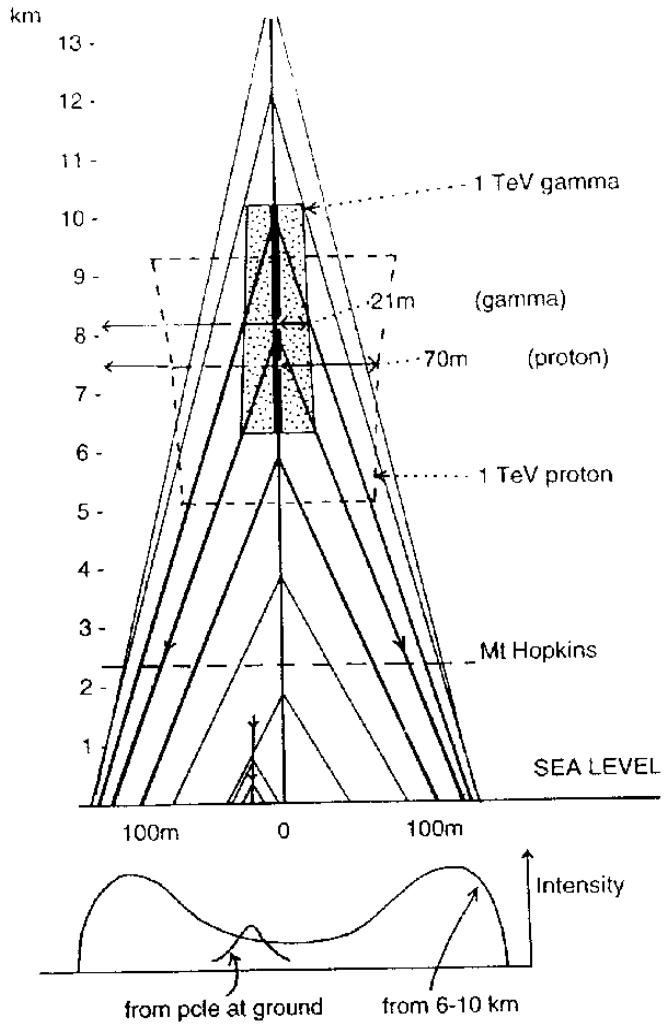


Figure 19: This figure sketches the Cherenkov light production of a 1 TeV photon. It illustrates how the opening angle increases with the density of the air. Most of the light is emitted between a height of 10 km and 6 km. (taken from [Hil96])

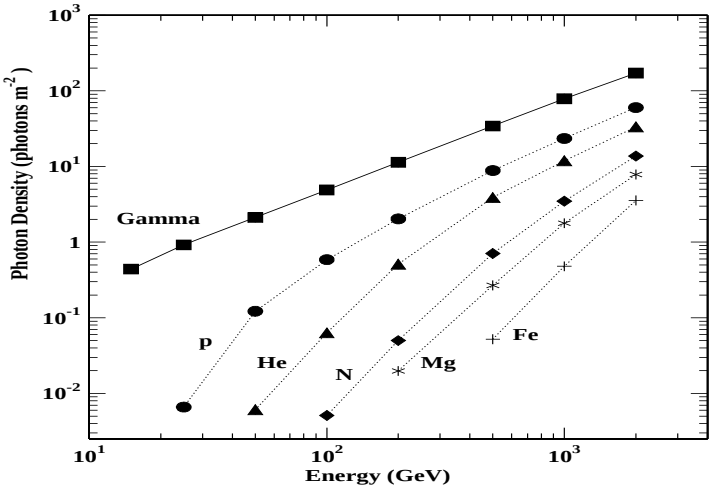


Figure 20: Cherenkov photon density (300 nm-600 nm) of different cosmic ray particles at 2000 m averaged over $50000\ m^2$ collection area (impact parameter $<\ 125\ m$), taken from [Ose01]

more in time scales of 20 minutes, depending on the humidity of the air. During moon light, the night sky background can be up to ***two orders of magnitude higher***.

The ***brevity*** of the light flash defines the properties that a camera must have: ***Single photon counting capability***, ***ultra short recording time*** of a few nano seconds, and the ability to ***withstand*** a rather high continuous photon background rate. This can be achieved with a camera consisting of pixels of photomultipliers, which are run at a sufficiently low gain to avoid damage by high currents, but still high enough to resolve at least a few photon pulses.

The size of the pixels should be smaller ***than 0.3***° to at least be able to resolve the lateral spread (To see the fluctuations in the lateral distribution, the pixel diameter should ideally be below 0.1°) because the main advantage of an imaging telescope is to be able to distinguish between background (hadronic showers) and signal (gammas showers) ***by seeing differences*** in the image of the showers.

To bolster the above mentioned numbers lets give a numerical example: For a 1 TeV gamma shower the maximum is located at about an altitude of 10 km and the interaction length is about 1 km at that air density. Assuming the telescope is located at 2.2 km elevation, then the interaction length corresponds to $\sim\ 0.1^\circ$ viewing angle in longitudinal direction for an impact parameter of p = 100 m. According to the NKG formula Equ. 1.26, the lateral spread for such a shower is half width half maximum d=65 m at s=1 (shower maximum) at that altitude of 10 km. A movement of 64 m perpendicular to the center-core connection line at the same impact parameter results in approximately 0.3° which corresponds to one CT1-pixel in lateral direction in the camera (see Fig. 1.14).

In the following sections I will give an overview of the HEGRA experiment and how the CT1 telescope, a major prototype for all Cherenkov telescopes has been implemented.

.2 The HEGRA experiment: An overview

The original HEGRA experiment (see Fig. 21) was proposed by the institute of physics the Universität Kiel and was built as a small scintillator array in 1988 on the Canary Island La Palma (28.75° N, 17.89° W). Gradually more institutes joined and the detector was enlarged. At present, the HEGRA collaboration consists of seven institutes: Universität Hamburg, Max-Planck-Insitut für Physik in Munich, Max-Planck-Institut für Kernphysik

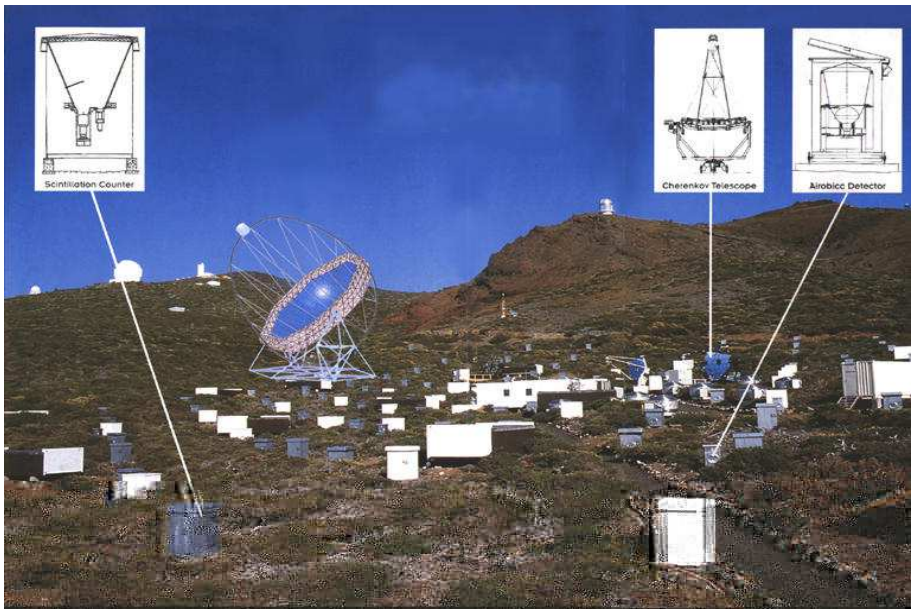


Figure 21: The HEGRA experimental site 1998. A collage shows the new MAGIC telescope which is currently under construction. On the right one can see the CT1 telescope which has been used to record the data for this thesis.

in Heidelberg, Universität Kiel, University of Madrid, Universität Wuppertal and the Yerevan Physics Institut in Armenia. The HEGRA experiment is located at the site of the ***Observatorio del Roque de los Muchachos*** at about 2200 m above sea level. In its 1997 setup, the experiment consisted of 17 Geiger counters, 244 scintillator counters, 77 wide angle Cherenkov counters and 6 Cherenkov telescopes. In the past year parts of the experiment have terminated and have been removed. At present only six Cherenkov telescopes are in use. Five of them are operated together in stereo mode in the so-called CT-system. The oldest telescope and original prototype of the experiment, is run in a stand-alone mode. This is the CT1 telescope. Only data from the CT1 telescope has been used for this thesis.

The CT1 telescope was installed 1992. In the beginning it had $5\,m^2$ of mirror area and a camera of 37 pixels with $\sim\ 3^\circ$ field of view (FOV) [Mir94]. In December 1994 the camera was replaced by a higher resolution camera with ***127 pixels*** and a $\sim\ 3^\circ$ FOV [Rau95]. CT1 was operated in this configuration until November 1997, at which time it was upgraded to $\sim\ 10\,m^2$ of mirror collection area made from aluminum mirrors. This represents the current setup used for this thesis. In this configuration ***it has a threshold of approximately 750 GeV***.

The second HEGRA Cherenkov telescope, CT2, was installed in 1993. In contrast to CT1, which has an equatorial mount, CT2 uses an alt-azimuth mount and is equipped with $8.5\,m^2$ mirror area. In the beginning CT2 had a camera with only 61 pixels, but it was upgraded in 1998 to a high resolution camera of 271 pixels and 4.9° field of view. CT2 was also operated in stand-alone mode until 1998, when it was included in the CT-system. In 1997 it had a threshold energy of approximately 1 TeV.

The Cherenkov telescope system, called ***CT-system***, consists of a set of four telescopes (namely CT2 to CT5). Together they are ***operated in stereo mode***. This means that all of the telescopes watch the same object and simultaneously record the same air shower. In this way the impact parameter and therefore the energy can be reconstructed much more precisely. The advantage of having several images of the same shower leads to a much ***improved gamma/hadron separation***. It is possible to obtain an ***almost perfectly clean data set of gamma showers*** when operating in a $\geq$ 3-fold telescope coincidence. The first

telescope of the CT system, CT3 was installed in 1995. CT4, CT5 and CT6 followed and were fully constructed end of 1996. All of these telescopes have a segmented mirror of $8.5\ m^2$ and a high resolution camera of 271 pixels and 4.9 $^\circ$ FOV. The energy threshold of the system is 500 GeV.

.3 The HEGRA CT1 telescope

The CT1 telescope has always been referred to as the '***prototype telescope***' because it was the first telescope of the HEGRA collaboration (see fig. 22) . It was constructed in order to gain experience with these types of telescopes and to see if Cherenkov telescopes, in general, are able to record weak astrophysical signals with sufficient signal to noise ratio to observe astronomical objects and to perform real physics. It has been one of the first imaging Cherenkov telescopes in the world. In contrast to most Cherenkov telescopes it has an equatorial mount instead of an alt-azimuth mount. This has both advantages and disadvantages. ***The advantages are***

- The operation of an equatorial mounted telescope is simpler because ***only one axis and one motor*** are necessary to rotate with ***constant speed*** in order to counteract the rotation of the earth.
- The ***coordinate system*** of the camera ***does not rotate*** with respect to the coordinate system of the sky. All of the bright stars in the field of view of the camera stay in the same position. This simplifies the Monte Carlo simulation and the analysis, as described later in chapter A.

The disadvantages are

- The axis of the telescope is ***never aligned perfectly*** to the earth's rotation axis and therefore the data needs a ***pointing correction***[3] later on.
- The construction ***requires a heavy counter weight*** for the mirror dish. This sets limitations on the accessible angular range. The use of counterweights requires a very stable and therefore ***expensive construction***.

Nowadays Cherenkov telescopes are constructed only with ***alt-azimuth mounts***.

.3.1 Technical details

In the present setup CT1 is equipped with segmented hexagonal aluminum mirrors of 490 cm focal length. A so-called Davis-Cotton configuration has been chosen in order to obtain an optimal shape of the reflector with respect to ***best images*** (least distortion). The total mirror area is $\sim\ 10\ m^2$ with a reflectivity better than 80%.

The camera consists of 127 10-stage EMI-9083A PMTs connected to hexagonal light concentrators, so-called Winston cones' which accept only light coming from a limited angular range in the direction to the mirrors in order to block stray-light and background light emerging from the side. The PMTs are operated at medium gain with only 8 stages coupled to fast preamplifiers to compensate for the reduced gain. This is done to avoid high anode currents generated by the light of the night sky (NSB) and bright stars in the field of view.

During dark nights (without moon light), the ***NSB*** gives approximately ***0.6 photon/(ns*pixel).*** This corresponds to approximately ***0.3 photoelectron/pixel*** (integrated over a gate length

[3]The word 'mispointing' is used in the sense that the center of the camera does not coincide exactly with the coordinates of the object to which the telescope is pointing to. A 'pointing correction' corrects the data for slight misalignments.

Figure 22: An image of the CT1 telescope in present state in La Palma on the Roque de los Muchachos. The data used in this thesis has been recorded with it.

of around 30 ns). This number increases by a factor of two with different weather conditions and by a ***factor of up to 20*** during moon light hours. In the latter case the high voltage (HV) of the PMTs is reduced to keep the current in an acceptable range. The diameter of one pixel is 0.25° ($\hat{=}$ 21mm). This is barely sufficient to resolve the lateral distribution of air showers. The camera is the most expensive part of an air Cherenkov telescope. Therefore one has to make compromises in the choice of the pixel-diameter and FOV. The total FOV is 3°. The maximum quantum efficiency (QE) of the PMT is 26% at 375 nm. The tracking error of the equatorial mount is 0.03° and the pointing error < 0.15°. The output of the pre-amplifiers is transfered to the counting container via coaxial cables.

.3.2 The trigger of CT1

The telescope triggers on incoming air-showers on the condition that within 6 ns (8-2 ns minimum overlap) two neighboring pixels out of any 127 show a signal higher than 50 mV, which corresponds to a superposition of approximately 13 photoelectrons in the PMTs. The trigger signal opens the gate of 127 charge sensitive (LeCroy) ADCs and initiates a readout signal. For each triggered event the direction of the telescope (via shaft encoder values) and a time stamp (via a rubidium clock) are recorded. The ***trigger rate*** depends on the ***HV*** of the PMTs, the ***zenith angle*** of the telescope and weather ***conditions*** on the mountain. The camera and the readout electronics are described in more detail in [Rau95]. For zenith angle $\vartheta = 0$ the trigger rate is approximately 5-8 Hz and the ***energy threshold is about 750 GeV.***

.4 The All Sky monitor of the RXTE satellite

Sources that emit high energy (HE) γ-radiation normally also show strong keV-MeV γ-emission. These measurements have been carried out by satellite borne instruments be-

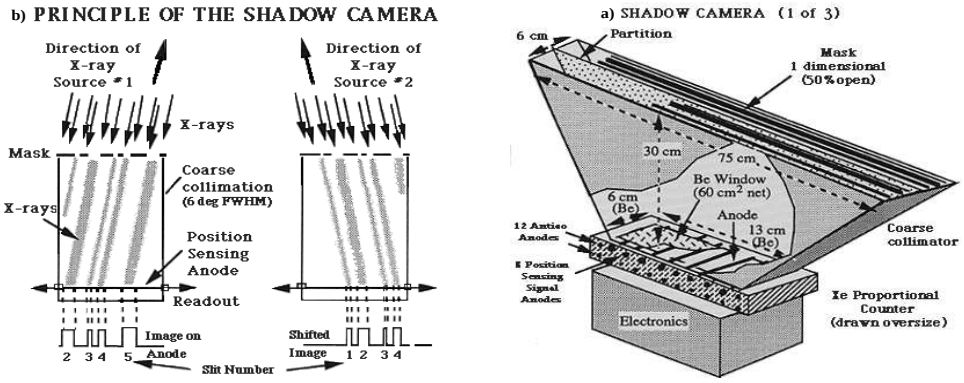


Figure 23: a) Principle of the Shadow Camera b) Schematic diagram of the Scanning Shadow Camera of the all sky monitor of the RXTE satellite.

cause the earth's atmosphere is an efficient shield. In chapter B, concerning the analysis of Mkn 421 observations, I will make use of x-ray observation data from the ***RXTE satellite***. Here I give a brief overview of this instrument.

The RXTE satellite [Bra93], launched on December 30, 1995, was designed to probe cosmic X-ray sources on variability time scales in the range of milliseconds to years. On board of the satellite three experiments have been installed which span an energy range of 2 keV to 250 keV. They are:

- ***Proportional Counter Array*** (***PCA***): The PCA consists of five xenon gas proportional counters with a collection area of about 6200 cm^2. The PCA is sensitive to X-rays in the energy range from 2 keV to 60 keV and can measure short term variability down to the microsecond level.

- ***High Energy X-ray Timing Experiment*** (***HEXTE***): HEXTE consists of two independent clusters of four NaI/CsI phoswich scintillation detectors. The energy range is 15-250 keV and the time resolution is again a few microseconds.

- ***All Sky Monitor*** (***ASM***): The ASM is the detector whose data has been used in this thesis. The ASM consists of three Scanning Shadow Cameras (SSCs) and covers an ***energy range from 2 to 12 keV***. The SSC has a slit mask at the entrance (see fig. 23), that produces a characteristic shadow pattern in the Position Sensitive Proportional Counters (PSPC) which are filled with Xenon. The idea is that each source in the sky produces such a pattern in the detector and that the pattern of all sources superpose. Deconvoluting the image delivers the intensity and direction of the individual sources. The ASM scans 80 % of the sky every 1.5 h and is used to measure long-term variability (hours to months) of bright X-ray sources due to the rotation of the satellite. All three SSCs point in different directions of the sky. ***The data is delivered in 90 second bins***.

For a more complete description of ASM and HXTE see [Lev96, Gru96].

Appendix A

The Analysis of shower images

In the first chapter, concerning theory and detectors, I gave a short overview of γ[1]-rays from their production in the source up to their development as air showers and their detection by Cherenkov telescopes. A much higher quantity of hadronic particles impinge onto the earth's atmosphere than ***γ-rays***, which themselves only account for ***less than 0.1%*** of all cosmic rays. One of the main challenges of γ-astronomy is therefore the separation of γ's from the hadronic background. A ***certain*** gamma/hadron separation is already applied by the ***telescope trigger***, which only responds to signals in a limited time window and takes advantage of the fact that Cherenkov photons from hadronic showers arrive with a ***larger time spread*** and have a lower light yield for the same energy as compared to γ's. From the moment that they have been recorded and measured in the form of Cherenkov images, the key issue becomes to ***separate*** gamma shower events from hadronic shower events as efficiently as possible.

The data used in this chapter are a part of the complete dataset of Mkn 421 from February to May 2001 and account for approximately ***167 hours*** of observation time (from a total of 250 hours that will be analyzed in the last chapter). It was recorded by the ***CT1*** telescope of the HEGRA collaboration on La Palma on the Roque de los Muchachos. The data has been preprocessed by a filter and the preproc-program written by Dirk Petry of the Max Planck Institut für Physik in Munich. The filter checks the data for some simple errors in mainly two ways: a) ***rejection of noise triggered events*** by applying a two next neighbor software-trigger and b) checks for ***correct positioning*** of the telescope. After the filter cuts, the trigger rate in zenith position is approximately 2.6 Hz to 3 Hz. Then, the preprocessing program which does the ***calibration*** is applied. It converts the signal into photoelectrons for each pixel and determines their pedestal RMS values from calibration runs. The next steps are done by a software package written in ***C++/ROOT*** which consists of about 50.000 lines of code that I developed. The software reads the calibrated data output and perform all the algorithms and analysis procedures discussed in this chapter.

For telescope efficiency studies, flux calculations and energy calibrations, Monte Carlo (MC) studies are essential. The MC simulation of air showers plus their imaging on the telescope have been done by Dorotha Sobczynska of the University of Lodz in Poland.

In the shower and telescope simulation by Dorotha Sobczynska the simulation of the night sky background (NSB) is not included and must be performed separately for each source analyzed. Therefore the ***simulation of the NSB*** is done within the program package developed for this work and will be described in the following sections.

This chapter is structured in the following way. After a quick description of the ***calibration procedure*** of preproc, an overview of the ***Monte-Carlo simulation*** is given. The next section deals with the ***classical image parameters*** and various ***separation methods*** (for ***instance***, ***static cuts***, ***dynamic cuts*** and a ***modified linear discriminant method*** (LDA))

[1] Note: Throughout this document the abbreviation 'γ' refers to a high energy photon (>1 GeV)

which are used to quantify the discriminating power of a given image parameter set. ***New image parameters*** are introduced and classified by using the ***LDA***. In the last section the ***estimation of energy***, ***pointing corrections***[2], ***unfolding of spectra, flux calculations*** and determination of the ***integrated light flux*** (above 1 TeV -> light curve) will be discussed.

A.1 The calibration of the telescope data

Here I wish to briefly discuss how the raw data from the CT1-telescope is calibrated. The calibration is done via a software 'preproc' which was written by Dirk Petry of the Max Planck Institut für Physik in Munich and significantly modified and improved by Martin Kestel, also in Munich. It directly reads the recorded rawdata of the telescope, which consists of ***pedestal runs***, ***calibration runs*** and ***observation runs***. The data simply consists of the ADC-values of each pixel that is recorded when a trigger occurs.

During pedestal runs random trigger images of the sky at the position of a source are taken. As the name implies, they are used to determine the pedestal position (in ADC counts), which means the zero line, and its RMS value of the pedestal peak. The RMS value is a measure of the light of the night sky and is proportional to the square-root of the PMT current. This will be explained in more detail in the next section concerning Monte Carlo simulation.

The calibration runs consist of a train of LED light-pulser events of equal amplitude. From the position of the signal peak, its RMS and the position of the pedestal, ***the number of photoelectrons*** can be calculated using the so-called ***excess noise factor method*** [Mir00, Sch01].

$$N_{PhE} = F\frac{(\mu - \mu_0)^2}{\sigma^2 - \sigma_0^2} \tag{A.1}$$

where μ and μ_0 are the signal and pedestal position and σ and σ_0 are the standard deviations of the signal peak and the pedestal peak. By assuming that N_{PhE} is equal for each pixel (so-called flat-fielding) and by knowing the average excess noise factor for all PMTs in the camera a ***conversion factor*** for ADC-channels->PhE for each pixel is calculated.

The precision of this method is estimated to be approximately 10 % [MirCom]. The conversion factors are calculated by evaluating the calibration and pedestal runs. Once the conversion factors are known they are used by the 'preproc' program to ***determine the number of PhE*** for each pixel of all the events in an observation run.

A.2 Monte Carlo Simulation of air showers

The Monte Carlo (MC) simulation of air showers and their imaging on the telescope camera is a very important part of the analysis, since it helps to understand the difference between gamma showers and hadronic showers. The simulation of the trigger of the telescope and the imaging of the shower onto the focal plane are absolutely mandatory to calculate the ***cut and trigger efficiencies***, which ultimately determine the ***effective areas*** after cuts and the flux. The ***energy calibration*** of the data is done by using MC data. This is done by first finding a good estimation of the shower energy of MC gamma showers and then applying this result to real data. This topic will be discussed later on.

The showers have been simulated by a base program from the Max Planck Institut in Karlsruhe/Germany called ***CORSICA*** which uses experimentally determined cross sections and complex atmospheric models to simulate the extended air showers by tracking each particle in the shower individually. As mentioned above, the simulation of the showers and especially the telescope reflector simulation have been performed by Dorotha

[2]The word 'mispointing' is used in the sense that the center of the camera does not coincide exactly with the coordinates of the object to which the telescope is pointing to. A 'pointing correction' corrects the data for slight misalignments.

Sobczynska. For the imaging onto the camera, the exact mirror and camera geometry have been taken into account, including optical imperfections. To obtain good agreement with the real measurement, the influence of the ***night sky background*** (NSB) and the light of starfields also need to be simulated. This is explained in the following section.

A.2.1 Simulation of the night sky light

In between to the normal data taking runs (for the observation of an astronomical object), ***calibration runs*** and ***pedestal run***s are also performed. The calibration runs are used to determine the ***conversion factor*** of ADC-channel/PhE. The pedestal runs consist of images of the night sky using random triggers to ensure the exclusion of shower light. The pedestal runs are used to measure the ***zero line*** of the ADC (pedestal) and its ***RM***S. The pedestal RMS of each pixel is calculated by taking the variance of the measured signal in each pixel. Both, the electronic noise and the light of the night sky (NSB) including starlight, contribute to the RMS. The NSB accounts to approximately 0.8-3 PhE per pixel (depending on observation conditions) within the time window of the ADC (approximately 30 ns).

To understand the ***shape*** of the resulting pedestal distribution, which we want to simulate, it is necessary to understand the camera electronics in significant detail. The light is recorded by PMTs which have a photocathode which converts photons into photoelectrons with a certain QE. The signal seen at the output of the PMT is the amplified signal of the PhEs hitting the first dynode. The number of PhE within the time window of the ADC is Poisson distributed. The further amplification has fluctuations due to its statistical nature. The dominant contribution comes from the first dynode. This ***additional noise is called excess noise*** (***F***). For PMTs it is defined as

$$F^2 = 1 + \frac{var(Single\,PhE\,Peak)}{mean^2(Single\,PhE\,Peak)} \tag{A.2}$$

The definition of F simply describes the increase of the noise of an incoming signal after amplification through the dynode system of a PMT. More generally (for any type of amplifier):

$$F^2 = \frac{Signal^2_{input} / \left(\sigma^2_{input} - \sigma^2_{el}\right)}{Signal^2_{output} / \left(\sigma^2_{output} - \sigma^2_{el}\right)} \tag{A.3}$$

The noise at the input and at the output is understood as noise without the electronic noise of the amplifier $\sigma^2 = \sigma^2_{real} - \sigma^2_{el}$. In the simulation the ***Poisson distributed signal*** of the PhE, with mean λ, has to be ***folded*** with a ***Gaussian distribution*** (coming from excess noise and electronic noise). The output signal of the PMT (the gain has been normalized to one) is then:

$$f_\lambda(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\,\lambda^n}{n!} \frac{e^{-\frac{(x-n)^2}{2\sigma_n^2}}}{\sigma_n\sqrt{2\pi}} \tag{A.4}$$

with a variance of

$$\sigma_n^2 = \quad n\left(F^2 - 1\right) + \sigma^2_{el} \tag{A.5}$$

σ_0^2 is the variance of the pedestal which is equal to the electronic noise contribution $\sigma_0^2 = \sigma^2_{el}$. A consistency crosscheck of the output distribution function $f_\lambda(x)$ to the definition of the excess noise factor can be found in Appendix A.

The variance of the electronic noise σ_{el} in the case of the CT1 camera electronics is estimated to be equivalent to 0.5 PhE. The excess noise is ***typically*** approximately $F^2 = 1.3$.

A very important point that has to be taken into account is that the output of the PMTs is coupled to the transimpedance amplifier via a ***capacitance*** such that only fast pulses are

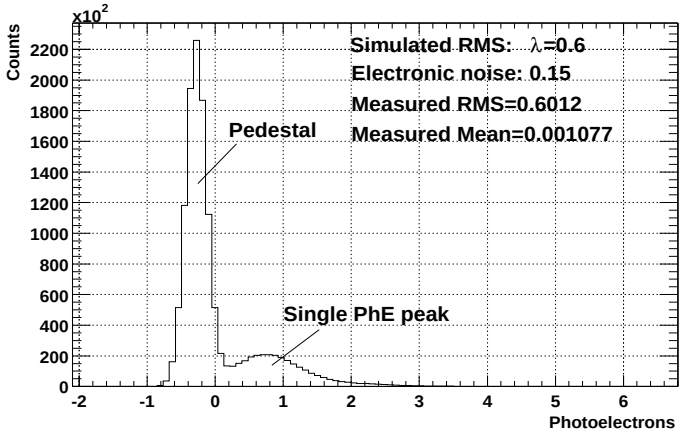


Figure A.1: This plot shows the simulation of the night sky background. Here the electronic noise was set to only 0.15 PhE in order to be able to see the single PhE peak.

amplified. ***Only the fluctuations*** of the DC like NSB are amplified. Therefore the average has to be subtracted:

$$f_{LONS}(x) = f_{\lambda}(x) - \lambda \tag{A.6}$$

$f_{LONS}(x)$ is the final simulated night sky background (NSB). Fig. A.1 shows a simulated NSB with a total RMS of 0.6 PhE. Here the electronic noise was set to only 0.15 PhE in order to be able to see the single PhE peak.

For subsequent data analysis an electronic noise of 0.5 PhE was used which corresponds to the estimated electronic noise of the CT1 camera read-out.

In order to treat the night sky background correctly, the pedestal distributions for each pixel are simulated on top of the MC shower images according to the measured pedestal RMS of the recorded data run. Since the starfield changes from object to object and the night sky background depends strongly on the zenith angle, the night sky background is simulated differently for each dataset. For the optimization of cuts or the training of the LDA it is very important to divide the recorded data into zenith angle bins and simulate the NSB ***exactly with the same zenith angle distribution*** on top of the MC shower images because the NSB depends on the zenith angle, which will be shown later. Within each zenith bin angle the available events are distributed in equal numbers among the recorded runs. Then for each run the pedestal information is taken and the NSB is simulated accordingly. By this procedure it has been ***ensured*** that the night sky background of the MC showers resembles as much as possible the NSB in the recorded data ***with the same zenith angle distribution***.

Finally, after our gamma data sample has been prepared to be comparable to the recorded data set and selection cuts have been applied, we are interested in determining the flux.

A.2.2 Trigger efficiencies, cut efficiencies and effective areas

The fluxes seen by the telescope have to be calculated out of the measured ***event rate*** and the ***efficiencies*** of the telescope. In general the ***differential flux*** is defined as

$$\frac{dF}{dE} = \frac{dN}{dE\,dA\,dt} \tag{A.7}$$

Where F is the flux, E is the energy, A is the area, N is the number of particles (γ's) and t is the time. The telescope ***efficiencies*** are defined as the ratio of the number of so-called

excess events N_{Excess} and the real existing number of events N_{Real}.

$$\varepsilon_{tot} = \frac{dN_{Excess}}{dN_{Real}} \tag{A.8}$$

The ***excess events*** $N_{Excess} = N_{Pass} - N_{Back}$ are the measured events by the telescope which passed the γ-selection cut minus the background events. The efficiencies of the telescope can only be obtained by ***MC simulation*** since we do not have a test beam with cosmic γ's. Thus, the ***efficiencies*** are defined as the ***ratio*** of the events that passed the γ-selection cut $N_{Selected}\,(E,\theta,r)$ and the total simulated events $N_{Simulated}\,(E,\theta,r)$.

$$\varepsilon_{Total}\,(E,\theta,r) = \frac{dN_{Selected}\,(E,\theta,r)}{dN_{Simulated}\,(E,\theta,r)} \tag{A.9}$$

The number of events that passed the selection cut (selection efficiencies) are ***dependent*** on the ***energy*** E, the ***zenith angle*** θ and the ***impact parameter*** r. The total efficiency and can be split into ***two parts***, the trigger efficiency of the telescope trigger electronics and the cut efficiency of the selection cut (in the data analysis):

$$\varepsilon_{Total}\,(E,\theta,r) = \varepsilon_{Trigger}\,(E,\theta,r) \cdot \varepsilon_{cut}\,(E,\theta,r) \tag{A.10}$$

The best way to ***introduce*** the ***concept of effective areas*** $A_{eff}\,(E,\theta,r)$ is by ***calculating*** the excess events out of the ***flux:***

$$\begin{aligned}
\frac{dN_{Excess}}{dE} &= \int \frac{dF}{dE}\varepsilon_{Total}\,(E,\theta,r)\;dt\,dA \qquad (A.11)\\
&= \frac{dF}{dE}\sum_{\theta_i} T_{obs}^{\theta_i}\int\int \varepsilon_{Total}\,(E,\theta_i,r)\;dr\,d\phi\\
&= \frac{dF}{dE}\sum_{\theta_i} T_{obs}^{\theta_i}\,A_{eff}\,(E,\theta_i)
\end{aligned}$$

The integral over the time has been transformed into a sum over zenith angle bins θ^i because the effective areas depend on the zenith angle and a binning of the recorded telescope data into zenith angle bins is unavoidable. The integration over the impact area has been separated into an integration over the radius r (rotational symmetry) and the azimuth angle ϕ.

The ***effective area*** is therefore the integral over the efficiencies (which are obtained from MC studies):

$$\begin{aligned}
A_{eff}\,(E_k,\theta_l) &= \int\int \varepsilon_{Total}\,(E_k,\theta_l,r)\;d\phi dr \qquad (A.12)\\
&= 2\pi\int_0^{\infty} \varepsilon_{Total}\,(E_k,\theta_l,r)\;dr\\
&= \pi\sum_{i=1}^{N}\frac{N_{Selected}\,(E_k,\theta_l,r_i)}{N_{Simulated}\,(E_k,\theta_l,r_i)}\left(r_{up}^2(i) - r_{low}^2(i)\right) \qquad (A.13)
\end{aligned}$$

The effective area ***depends*** on the ***energy*** and the ***zenith angle*** (see Fig. A.2). $r_{up}\,(i)$ and $r_{low}\,(i)$ are the upper and lower edge of the impact parameter bin i, respectively. A ***binning*** in impact parameter, energy and zenith angle is ***unavoidable.*** As above $N_{Selected}\,(E_k,\theta_l,r_i)$ is the number of events in the according energy/zenith angle/impact parameter-bin that passed the selection cut and $N_{Simulated}\,(E_k,\theta_l,r_i)$ is the number of events that have been simulated originally for the same bin.

The effective areas are also slightly dependent on the shape of the flux spectrum because of the binning in energy bins. In order to avoid systematic errors, each MC event is weighted with $W\,(E)$ in such a way that the weighted MC distribution forms the desired spectrum. The desired spectrum should be similar to the spectrum of the source that is measured. This can be achieved by an iterative process for which the measured spectrum is placed back into the effective area calculation. The systematic error here is reduced when taking small energy bins. The weights are normalized to one.

The ***effective area*** is the conversion factor from the ***excess event rate*** to ***gamma ray flux*** which is calculated in the following way:

$$\frac{dF}{dE} \;=\; \frac{\frac{dN_{Excess}}{dE}}{\sum_{\theta_i} T_{obs}^{\theta_i}\, A_{eff}\,(E,\theta_i)} \tag{A.14}$$

The errors on the fluxes are calculated via Gaussian error propagation from errors on N_{excess} and A_{eff}. The errors on the effective areas are statistical Poissonian fluctuations from $N^{i}_{selected}$ and from $N_{Simulated}\,(E,\theta,r_i)$ and they add up quadratically:

$$\sigma^2_{A_{eff}} = \pi \sum_{i=1}^{N} \left[\frac{\sqrt{N_{selected}\,(E,\theta,r_i)}}{N_{Simulated}\,(E,\theta,r_i)} \left(r^2_{up}(i) - r^2_{low}(i)\right) \right]^2 \tag{A.15}$$

A.3 Gamma/hadron separation methods

For this analysis I used an approach which first describes the shower image by using ***so-called image parameters***. The distributions of the image parameters for γ's and hadrons exhibit some differences. A cut method which uses these differences separates the two event types using image parameters was developed. Another approach, which has not been used in this thesis, is to directly apply a ***maximum-likelihood fit*** to the shower images themselves [CAT89]. This approach depends on how well it is possible to model the shape of gamma showers and hadronic showers which need to be used as templates for the fit.

In this section I wish to introduce as a first step the classical image parameters, called '***Hillas parameters***' which are named after A. Hillas who invented them in 1985 [Hil85]. Afterwards, simple static cuts and a linear discriminant analysis (LDA) will be described. Eventually the LDA will be used mainly as a tool to quantify the discrimination power of different image parameter sets as it increases in discrimination power of a given parameter set, compared to the static cuts.

A.3.1 The classical Hillas parameters to describe shower images

As shown in chapter 1, concerning theory and detectors, the main differences between ***gamma photon initiated*** showers and ***hadron initiated showers*** lie in their ***geometrical*** structure and, to a lesser extend in their time structure (which is not discussed in this thesis). Due to the longer interaction length of the hadronic interaction, the ***hadronic showers*** show ***more fluctuations*** in their image than do the electromagnetic showers. In addition, hadronic showers have a ***wider lateral distribution***. Therefore, A. Hillas therefore proposed to calculate the geometrical variances of the shower image which are different for hadronic and gamma showers. He introduced the first and second moments of the shower image. In the following, I describe a form with ***generalized weights*** w_i which will be used later on (where x_i and y_i are the coordinates of pixel i). In case of the classical Hillas parameters, the weight w_i is the charge collected by pixel i in photoelectrons $w_i = q_i$. Fig. A.3 illustrates the geometric meaning of the image parameters.

The ***first moments*** are

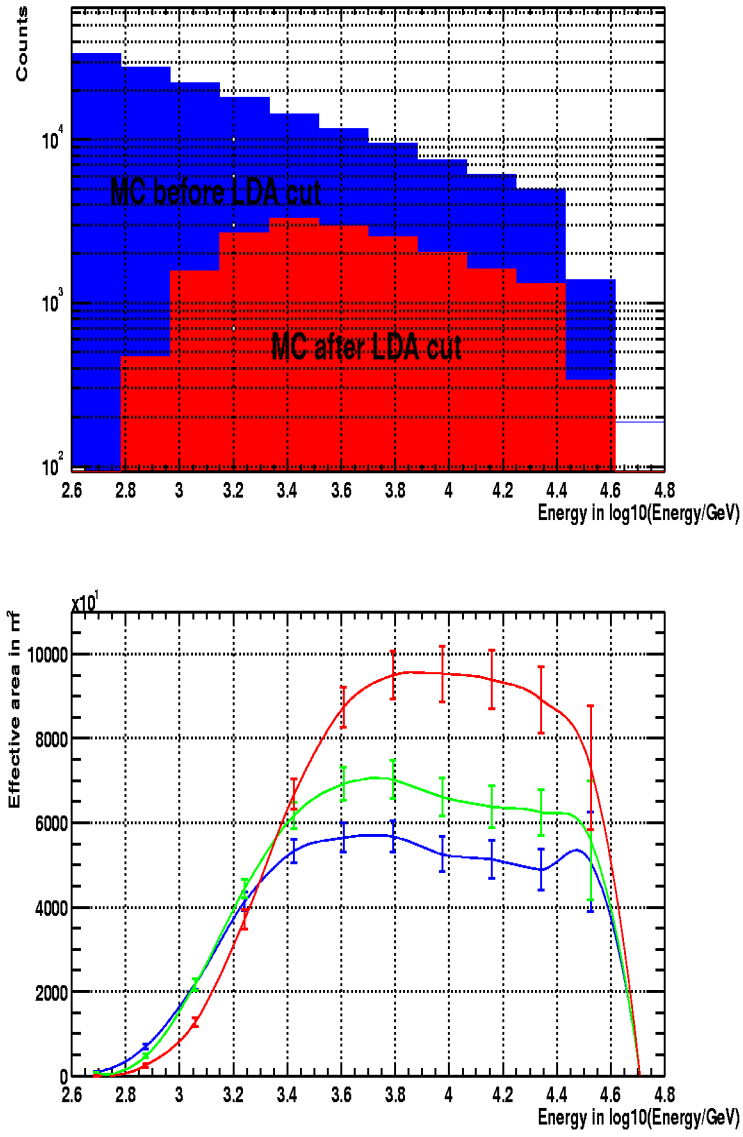


Figure A.2: The upper plot shows the ***total simulated MC event distribution*** (blue histogram) and the ***events that passed the selection cut*** (red curve), integrated over all impact parameters. The lower plot shows the ***effective areas*** for the selection cuts developed in this thesis using a power-law spectrum with spectral index $\alpha = -2.8$. The three curves represent the effective areas for ***three*** zenith angles: 12° (blue) , 32° (green) and 50° (red).

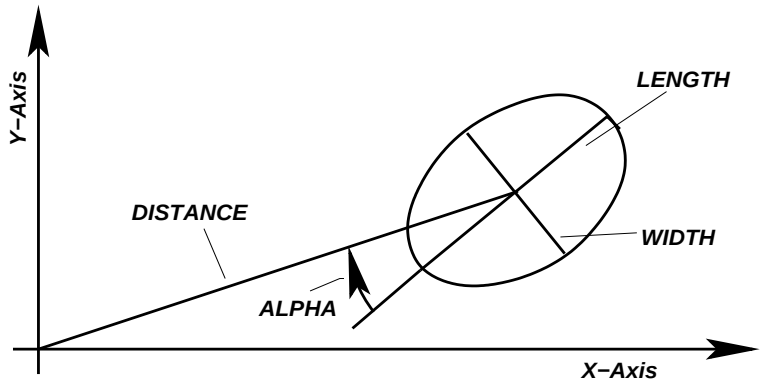


Figure A.3: An illustration of the geometric meaning of the image parameters.

$$\begin{aligned} \langle x \rangle &= \frac{\sum_{i=1}^{N} x_i w_i}{\sum_{i=1}^{N} w_i} \\ \langle y \rangle &= \frac{\sum_{i=1}^{N} y_i w_i}{\sum_{i=1}^{N} w_i} \end{aligned} \tag{A.16}$$

and the ***second moments***

$$\begin{aligned} \langle x^2 \rangle &= \frac{\sum_{i=1}^{N} x_i^2 w_i}{\sum_{i=1}^{N} w_i} \\ \langle y^2 \rangle &= \frac{\sum_{i=1}^{N} y_i^2 w_i}{\sum_{i=1}^{N} w_i} \\ \langle xy \rangle &= \frac{\sum_{i=1}^{N} x_i y_i w_i}{\sum_{i=1}^{N} w_i} \end{aligned} \tag{A.17}$$

The ***variances*** can be calculated in the classical way:

$$\begin{aligned} var(x^2) &= \langle x^2 \rangle - \langle x \rangle^2 \\ var(y^2) &= \langle y^2 \rangle - \langle y \rangle^2 \\ covar(xy) &= \langle xy \rangle - \langle x \rangle \langle y \rangle \end{aligned} \tag{A.18}$$

The variances can be put together to form the so-called ***covariance matrix***

$$M = \begin{pmatrix} var(x^2) & covar(xy) \\ covar(xy) & var(x^2) \end{pmatrix} \tag{A.19}$$

The covariance matrix can be used to describe a ***two dimensional Gaussian distribution*** function

$$\begin{aligned} G(x,y) &= \frac{1}{\sqrt{(2\pi)^2 \det M}} e^{-\frac{1}{2}\mathbf{u}M^{-1}\mathbf{u}} \\ \mathbf{u}M^{-1}\mathbf{u}\left(1-\rho^2\right) &= \frac{u_1^2}{var(x^2)} - 2\rho\frac{u_1 u_2}{\sqrt{var(x^2)\,var(y^2)}} + \frac{u_2^2}{var(y^2)} \end{aligned} \tag{A.20}$$

with the ***correlation coefficient***

$$\rho = \frac{covar(xy)}{\sqrt{var(x^2)\,var(y^2)}} \tag{A.21}$$

and

$$\begin{aligned} u_1 &= x - \langle x \rangle \\ u_2 &= y - \langle y \rangle \end{aligned} \tag{A.22}$$

The two dimensional Gaussian distribution is ***rotated*** in the camera coordinate system by an angle given by

$$\tan 2\phi = 2\frac{covar(xy)}{var(x^2) - var(y^2)} \tag{A.23}$$

By ***diagonalizing the matrix*** M, one obtains the longitudinal (LENGTH) and the lateral (WIDTH) variances of the shower image. By introducing the helper variables

$$\begin{aligned} d &= var(y^2) - var(x^2) \\ z &= \sqrt{d^2 + 4covar^2(xy)} \\ u &= 1 + \frac{d}{z} \\ v &= 2 - u \\ w &= \langle x \rangle^2 \langle y^2 \rangle - 2 \langle x \rangle \langle y \rangle \langle xy \rangle + \langle y \rangle^2 \langle x^2 \rangle \end{aligned} \tag{A.24}$$

WIDTH and ***LENGTH*** can be defined can be defined as:

$$WIDTH = \sqrt{\frac{var(x^2) + var(y^2) - z}{2}} \tag{A.25}$$

$$LENGTH = \sqrt{\frac{var(x^2) + var(y^2) + z}{2}} \tag{A.26}$$

Additional ***useful image parameters*** include:

$$DIST = \sqrt{\langle x \rangle^2 + \langle y \rangle^2} \tag{A.27}$$

$$AZWIDTH = \frac{\sqrt{w}}{DIST} \tag{A.28}$$

$$\begin{aligned} MISS &= 0.5\left(u\langle x\rangle^2 + v\langle y\rangle^2\right) - \frac{2covar(xy)\langle x\rangle\langle y\rangle}{z} \\ ALPHA &= \arcsin\frac{MISS}{DIST} \\ m_{ax} &= \frac{d + \sqrt{d^2 + 4covar(xy)^2}}{2covar(xy)} \\ m_{ce} &= \frac{\langle x \rangle}{\langle y \rangle} \end{aligned} \tag{A.29}$$

$$pr = m_{ax} m_{ce} \tag{A.30}$$

$$SIGN = \begin{cases} 1\,if\,(pr \geq 0)\,\&\,(m_{ax} \geq m_{cer}) \\ -1\,if\,(pr \geq 0)\,\&\,(m_{ax} < m_{ce}) \\ 1\,if\,(pr < 0)\,\&\,(m_{ax} \leq \frac{1}{m_{ce}}) \\ -1\,if\,(pr < 0)\,\&\,(m_{ax} > \frac{1}{m_{ce}}) \end{cases} \tag{A.31}$$

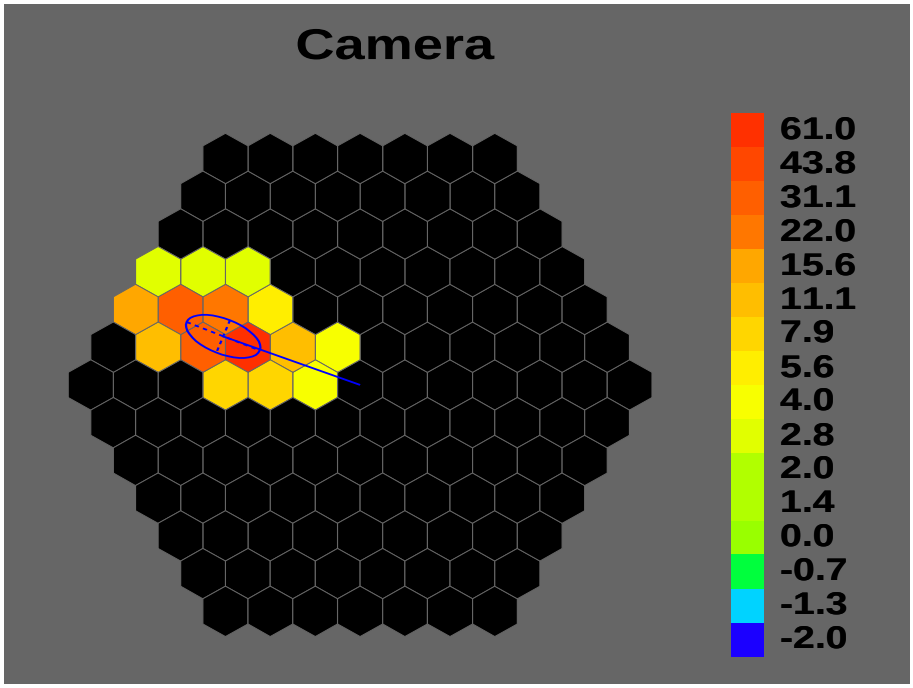


Figure A.4: The CT1 camera with a simulated gamma shower with its Hillas ellipse after image cleaning.

$$SIZE = \sum_{i=1}^{N} q_i \tag{A.32}$$

$$CONC = \frac{q_{1.} + q_{2.}}{SIZE} \tag{A.33}$$

where ***DIST*** (Equ. A.27) is the distance from the camera center to the center of gravity of the shower in degrees and it depends on the energy, the impact parameter and the zenith angle of the γ (see section 1.4.1 and Fig. A.3).

AZWIDTH (Equ. A.28) has not been used in this thesis because it does not improve the discrimination power anymore.

ALPHA (Equ. A.29) is the most important cut parameter for point sources. It is defined as the clockwise angle from the longitudinal axis of the shower to the connection line between the center of the camera and the shower center of gravity. Since the center of the camera of the telescope usually points towards the point source itself, the shower axis of gamma showers points toward the center. Conversely, hadronic showers arrive from all directions and the distribution is flat for all angles for an infinitely large camera (for a finite size camera the ALPHA distribution is no longer uniform). Equ. A.29, shown above, gives values between 0° and 90°. ALPHA defines an angle between 0° and 180° or equivalently, from -90° to +90°. For some calculations ***the sign of ALPHA is needed***. It can be obtained from Equ. A.31.

Another very basic parameter is ***SIZE*** (Equ. A.32). This is simply the total charge collected in units of photoelectrons (PhE). It is the ***main estimator for the energy***, but depends also strongly on the zenith angle of the source and the impact parameter.

CONC (Equ. A.33) is the ratio of the sum of the two highest pixel charges to the total charge. Gamma showers have a smaller lateral and longitudinal electron distribution and peak more in the center of the image, in contrast to hadronic showers.

Therefore, ***CONC*** is bigger and ***WIDTH*** and ***LENGTH*** are ***smaller*** for gamma showers than for hadronic showers (see Fig. A.5). Since the camera of CT1 has rather

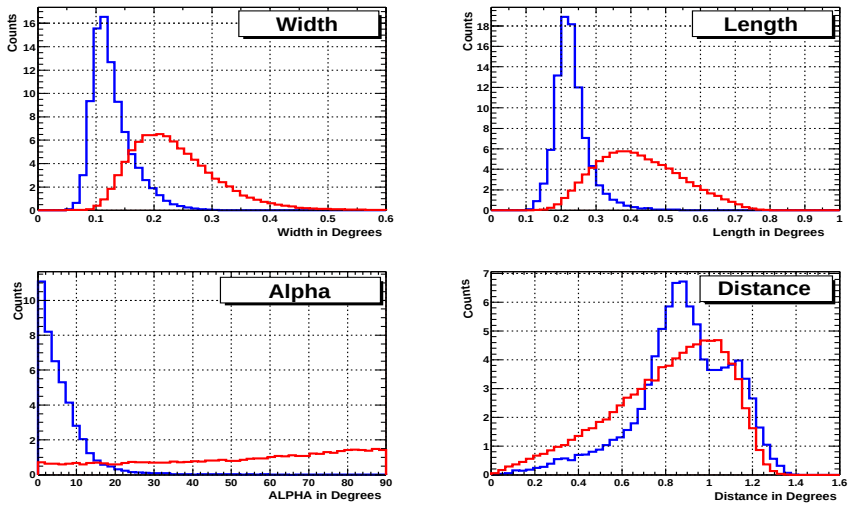


Figure A.5: The distributions of the four most important image parameters: WIDTH, LENGTH, ALPHA and DISTANCE . MC gammas are blue and recorded off-data is red.

large pixels (0.25° diameter), CONC is not very smoothly defined and shows a rather large spread. The following two figures show a) a ***typical gamma shower image*** with applied tail cut (see section about image cleaning) in the CT1 camera (Fig. A.4) and b) the ***distributions*** of the ***four most important*** image parameters (Fig. A.5). The second hump in the DIST distribution results from shower images that where larger than the radius of the camera and were therefore cut at the camera border (see leakage parameter). A selection cut of SIZE > 100 photoelectrons (PhE) has been applied.

Expression A.20 can be used to define a so-called ***covariance ellipse***, a line of equal probability, of 1 sigma (defined by WIDTH, LENGTH, DIST and ALPHA), around the center of weight. It is the so-called ***Hillas ellipse***. It is a description of the shower image (see Fig. A.4).

The definition of the image variances given above assumes that there is no noise background in the image. Unfortunately, this is never the case because we have NSB as we know. In Equ. A.17 it can be seen that the distance of the pixels enters as a square in the sum and, consequently, the calculation of the variances of the shower image gives entirely incorrect values. Therefore, an image cleaning (sometimes called a tail-cut), becomes necessary to remove, at least in part, the effect of the NSB.

A.3.2 Image cleaning algorithms to remove the night sky background in the camera

The image cleaning (or tail cut) ***removes*** the night sky background (NSB) in the camera. Without this it would not be possible to calculate the image variances, as explained above. Traditionally a dynamical image cleaning has been used [Pet97]. The RMS values of the pixels are measured in independent pedestal runs and the signal of the pixels are compared to their RMS values in order to decide if the pixels should be included in the parameter calculation or not.

In this analysis ***three different algorithms*** have been tested. In all three cleaning methods two cleaning parameters were introduced: The '***image core limit***' κ_c and the '***image border limit***' κ_s.

The 'classical' cleaning method

1. All pixels that have a signal ***larger*** than κ_c (image core limit) times the Pedestal RMS ***are retained.***

2. All pixels that have a signal ***larger*** then κ_s (image border limit) times the Pedestal RMS are retained ***if*** they have a neighbour with more than κ_c sigma.

3. Single pixels are ***removed*** if the signal is ***smaller*** than 5 times the Pedestal RMS.

The 'island' cleaning method

First, the 'classic' cleaning algorithm is applied. Afterwards, the image is being analyzed for ***islands*** (islands are isolated clusters of pixels which remain after a classical cleaning. The meaning of 'island' is illustrated in Fig. A.24.). Only the ***largest*** island (in terms of charge; this is the main shower image) ***is retained***, all the others are removed. The island finding algorithm is explained in detail in the section concerning mountains and islands.

The 'mountain' cleaning method

The method is based on the ***quantification of fluctuations*** in the shower image, which are different for hadronic showers and electromagnetic showers.

In the first step, the 'classical' image cleaning is again applied. Then a complex image structure analysis is performed with respect to the ***'mountain' structure*** (The meaning of 'mountain' is illustrated in Fig. A.24.) of the image. This will be described in detail in section A.4 where new image parameters are introduced. In a recursive procedure the image is divided into mountains. (The analogy of mountains or valleys is with respect to their signal content). The ***cut lines are the valleys*** between the mountains. The charge in each cluster (mountain) is summed up and the ***largest mountain is retained*** while the others are cleaned away.

Later, these three algorithms will be compared with each other, with respect to their capability of delivering the best image parameters.

General comments

Dynamical[3] image cleaning procedures have the advantage of ***retaining as much as possible*** of the image in order to have the ***largest*** amount of information about the shower available. On the contrary, 'dynamical' cleaning procedures also have some ***disadvantages*** compared to fixed level image cleaning methods. When the tail-cut becomes ***dependent*** on the pedestal RMS, which is actually a measure of the NSB and the starlight, then as a consequence, the values for WIDTH and LENGTH ***also*** become ***dependent*** on the NSB. In other words, WIDTH and LENGTH will change at different night sky and weather conditions and dependencies that have not been there before are artificially introduced into the analysis.

A.3.3 Static cuts and dynamical cuts: Dependence of the Hillas parameters on the energy, the zenith angle, the impact parameter and the night sky background

Having introduced the parameters that describe the shower image we now wish to use them for hadronic background ***suppression***. In the following section the image parameters have been calculated in the classical way with weights $w_i = q_i$.

[3]The expression 'dynamical' is understood in the sense that the cleaning level is chosen as a function of the NSB (pedestal RMS).

Static cuts and the optimization of cut intervals

The simplest method for gamma/hadron separation consists of a ***static cut*** on the image parameter values. These cuts are commonly called ***super cuts*** (Wipple collaboration, Ref. [Rey93]). The following values have been obtained by following the procedure described below and using a data set of Mkn 421 (flares Feb 8/9, 2001):∘

$$\begin{aligned} 0.05^\circ &\leq WIDTH \leq & 0.11^\circ \\ 0.1^\circ &\leq LENGTH \leq & 0.42^\circ \\ 0^\circ &\leq ALPHA \leq & 12^\circ \\ 0.31^\circ &\leq CONC \leq & 0.7^\circ \\ 0.5^\circ &\leq DIST \leq & 1.0^\circ \end{aligned} \tag{A.34}$$

The ***cut on ALPHA*** is a geometrical cut and only works for ***point sources***. The lower cut on DIST is applied because images with too small impact parameters yield showers that are too round and too close to the center of the camera (bad definition of ALPHA) which makes a discrimination impossible. The upper limit is used to remove showers that are excessively affected by a limited camera size and that are truncated at the border.

The ***cut efficiencies*** for γ's: $\varepsilon_G = N_C^G/N_{tot}^G$ and hadrons: $\varepsilon_H = N_C^H/N_{tot}^H$ describe the percentage of events that have been selected after application of the cuts on simulated MC events. N_{tot}^H and N_{tot}^G are the number of all triggered γ-events and hadron-events. They have been simulated with a certain spectral index (power law, spectral index α=1.5) . The quantity ε_G should be as large as possible, usually at least 50 %, and ε_H should be as small as possible. Usually it is 0.5 % (Background reduction factor 200). The ***quality factor*** is defined by

$$Q = \frac{\varepsilon_G}{\sqrt{\varepsilon_H}} \tag{A.35}$$

It is a measure of how well the background has been ***suppressed*** by keeping enough signal events. Typical values for the static cut are about ***seven***.

The algorithm for ***cut interval optimization*** changes the intervals in small systematic steps in order to maximize either the ***quality*** factor or the ***significance*** (applied to pure MC-samples). The significance is the signal to noise ratio:

$$\begin{aligned} S &= \frac{N_{on} - N_{off}}{\sqrt{N_{on} + N_{off}}} \\ &= \frac{N_C^G}{\sqrt{N_C^G + 2N_C^H}} \end{aligned} \tag{A.36}$$

where $N_{on} = N_C^G + N_C^H$ and $N_{off} = N_C^H$ for Monte Carlo samples. S converges to

$$S \to Q\sqrt{N_{tot}^G} \tag{A.37}$$

for

$$N_C^G \gg N_C^H \tag{A.38}$$

Unfortunately, the quantity N_C^G is usually rather of the same order as N_C^H which means Equ. A.38 is not fulfilled and as a consequence, the optimization of the quality will ***not*** optimize the significance. That implies that it is ***more appropriate to optimize on the significance*** Equ. A.36, rather than on the quality Equ. A.35.

Further, if the significance is used for cut optimization then ratio N_{tot}^G/N_{tot}^H (if applied to MC samples) has to have the ***same value*** of approximately 0.005 which is observed in nature.

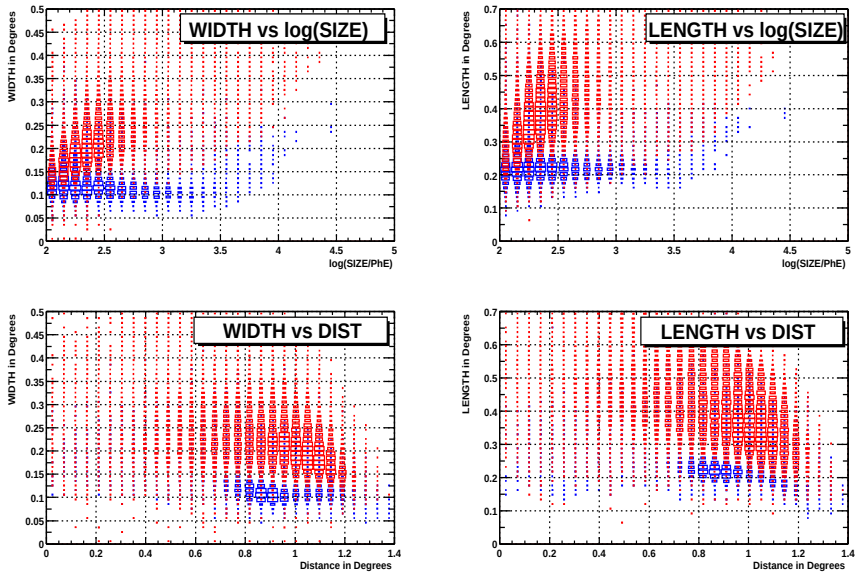


Figure A.6: The dependence of the parameters WIDTH and LENGTH ***on*** SIZE and DIST. MC gamma distribution is blue and recorded off-data is red.

The significance as defined above Equ. A.36 is ***not Gaussian distributed*** and is therefore not appropriate to estimate the ***probability***. This is important if one would like to claim a discovery with a small significance. An expression for the significance with Gaussian distribution has been defined by [LiMa83] as a log-likelihood ratio.

$$S = \sqrt{2}\left\{ N_{on} \ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{on}}{N_{on+N_{off}}}\right)\right] + N_{off} \ln\left[(1+\alpha)\left(\frac{N_{off}}{N_{on}+N_{off}}\right)\right]\right\}^{\frac{1}{2}} \quad \text{(A.39)}$$

where $\alpha = \frac{T_{ON}}{T_{OFF}}$ is the ratio of the observation times of the off-data sample and the on-data sample. This form of calculating the significance has been used ***in this analysis***.

By experience it has been found that the optimization on the quantity

$$P = S \cdot \sqrt{N_{ON} - N_{OFF}} \quad \text{(A.40)}$$

yields good results, because it forces the optimization algorithm to not only maximize the significance but to also keep a reasonable number of excess events $N_{Ex} = N_{ON} - N_{OFF}$.

The maximization procedure is nontrivial because each variation of a cut changes the sample and also because the data distribution has statistical fluctuations. The algorithm has to be able to handle these obstacles. The results of most algorithms ***depend on the initial value*** and are ***not reproducible***, only within a certain range and error. The result also varies slightly with small changes in the training data sample. This also pertains to dynamical cuts which are described in the following paragraph.

The optimization of dynamical cuts

The values of WIDTH, LENGTH, SIZE, DIST and CONC depend not only on the ***night sky background*** (expressed in pedestal RMS values) but also on the ***energy*** (estimator SIZE), the ***impact parameter*** (estimator DIST) and the ***zenith angle*** of the object in the sky (which is known and measured during the run).

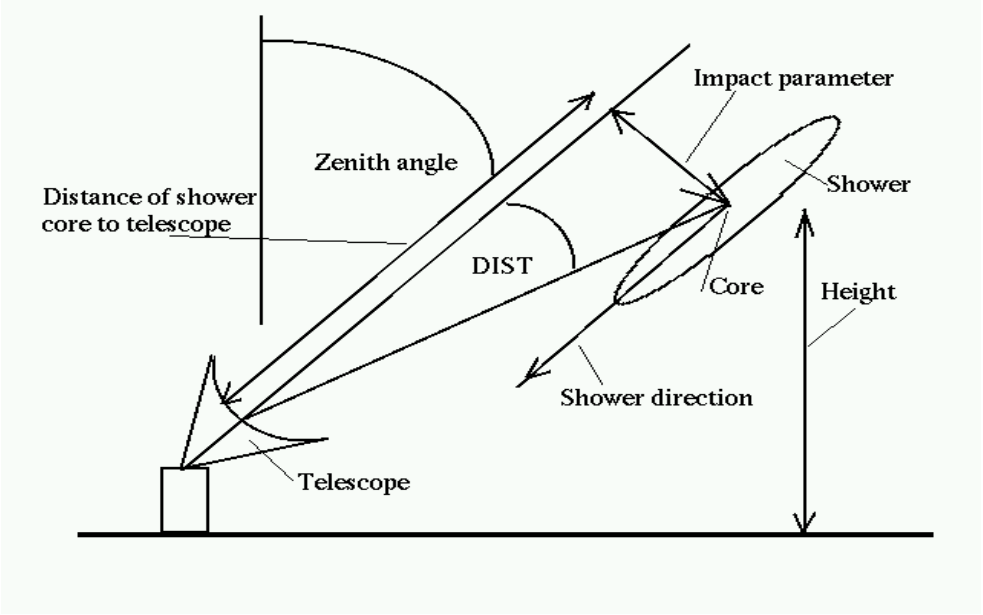


Figure A.7: This figure illustrates the imaging of a shower onto the telescope camera. It clarifies how a changing shower core position (changing energy and zenith angle) and a changing impact parameter effects the shape of the image.
It should be noted that Cherenkov light is not isotropy emitted but rather is only in the forward direction in light cones whose opening angle changes with the density of the air and the velocity of the particle.

The ***height*** of the shower maximum from the ground depends on the ***energy*** of the primary particle. The larger the energy, the deeper the shower develops in the atmosphere (see chapter 1, concerning theory and detectors). This changes the view of the shower and the image seems ***bigger*** and with a larger DIST, as illustrated in Fig. A.7.

A variation in the ***zenith angle*** has a similar effect. If this angle changes, the ***distance*** of the shower maximum to the telescope also changes. Tor ***larger*** zenith angles the maximum moves ***away*** from the telescope and the image seems generally smaller (smaller SIZE, WIDTH and LENGTH).

Similarly, a different ***impact parameter*** for the shower changes the ***ellipticity*** of the image.

Fig. A.6 shows the dependence of WIDTH and LENGTH on SIZE and DIST. It can be seen that for ***higher*** energies, gammas and protons are ***better separated,*** due to more available pixel information. In the figure a selection cut of SIZE > 100 PhE and a zenith angle < 30° has been applied.

To take ***into account the dependence*** just mentioned, the ***static cut intervals*** of Equ. A.34 need to be ***replaced*** by variable ones. This means that the ***cut intervals*** become ***functions*** of SIZE, DISTANCE and the zenith angle. This has been examined by Daniel Kranich in his PhD thesis [Kra01]. The dependence of WIDTH and LENGTH to the NSB can be corrected, as has shown Martin Kestel [Kes01].

Now let us assume that the cuts (static or dynamic) have been optimized and that we wish to apply them on a real dataset in order to extract the signal that we are interested in. It is clear that a pure signal dataset can ***never*** be achieved. The background can be suppressed ***to some degree*** but there will always be some background remaining together with the signal events. Apart from the fact that many hadronic showers look very similar to γ-showers just for statistical reasons, there exists also a certain fraction of hadronic showers that produce already at the first interaction a large fraction of energetic π^0 such

that these showers develop almost as pure electromagnetic showers. These showers are indistinguishable from the signal.

Optimization of cuts using real recorded data instead of Monte Carlo samples

It is possible to optimize cut values by only using ***recorded*** data and without touching MC data. The essential point is to ***know*** the background in the dataset, either by OFF-data or by some other kind of background estimation.

The procedure goes as follows:One can make a ***good guess*** for the initial values of the cut intervals, apply it to a set of real ON-data and estimate its background content after cut. Now the cut intervals can be optimized by ***maximizing the significance***.

Then the optimized cuts can be used on orthogonal datasets for analysis. This technique has been used by Daniel Kranich [Kra01].

A.3.4 Background determination by using the ALPHA-plot

Earlier it was mentioned that it is never possible to obtain a pure γ-events sample from a recorded data set because it will always be ***contaminated*** with hadronic background. For the calculation of the signal amplitude and its significance the background needs to be subtracted from the contaminated dataset. Therfore, a precise information about the background is needed. There are basically ***two possibilities***:

1. ***Two separate*** telescope ***runs*** are taken, one with the telescope pointing to a source (***ON-data run***) and the other one with the telescope pointing a few degrees away from the source (***OFF-data run***). After the selection cut are applied to both runs, the background can be determined precisely from the OFF data run and can be subtracted from the ON data run. Using Equ. A.39 the significance of the detection can be calculated.

2. For point source studies, the ***ALPHA plot*** can be used for this purpose since (see Fig. A.8) the γ events accumulate in the region up to about 12°, owing that the images point towards the center of the camera. By ***interpolating the shape of the background*** (which should be as flat as possible) the background in the region up to 12° can be estimated.

The latter method works quite well and was originally worked by G. Hermann in Heidelberg and has been successfully applied by Daniel Kranich [Kra01] in his thesis. The background estimation is achieved by fitting a ***simple polynomial*** to the distribution. The required conditions for the polynomial is that it has a ***horizontal tangent*** at 0° (because of the symmetry of the signed ALPHA distribution to the y-axis). While Daniel Kranich [Kra01] used a polynomial of second order, in this analysis a polynomial of order 2.5 is applied because background shape studies have shown that this is the function that provides the best fit to the background distribution.

$$f(ALPHA) = a + b\,|ALPHA|^{2.5};\ ALPHA \geq 0 \tag{A.41}$$

Assuming that the signal distribution of ALPHA is to first order approximated by a ***Gaussian distribution***, it can be used to estimate the signal events [Kra01]:

$$g(ALPHA) = \frac{N_{Signal}}{\sqrt{2\pi\sigma}} e^{-\frac{ALPHA^2}{2\sigma^2}} \tag{A.42}$$

The ***combination*** of both $h(\alpha) = f(\alpha) + g(\alpha)$ has 4 free parameters and gives a very good measure of the excess events by estimating the background from the ALPHA distribution in the region [20°, 70°] (see Fig.A.8).

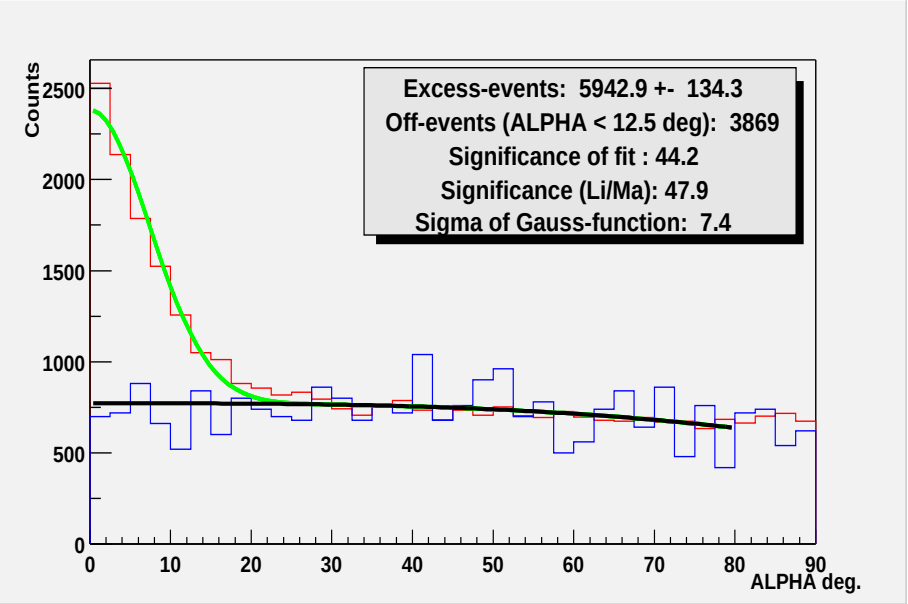


Figure A.8: This plot shows the data of a part of the dataset of Mkn 421 during its flare period from February to May 2001 (red line). The simple static cuts of Equ. A.34 have been applied. Background and excess events are estimated by fitting a combination of a Gaussian distribution (green line) and a polynomial (black line). For comparison, the blue line shows an off data dataset which has much lower statistics. We will see that the linear discriminant method and new image parameters improve the separation significantly. The new method finds almost a factor 3 more excess-events and it also doubles the significance. The expression 'significance of the fit' denotes the ratio of excess events and its error, both of them obtained by the fit. The expression 'significance (Li/Ma)' is the significance that has been calculated by using Equ. A.39 (with $\alpha = 1$).

This procedure works well if the ALPHA background distribution is ***quite flat***, which unfortunately is not necessarily the case since the limited diameter of the camera ***truncates*** events that are large or have sizable impact parameters. Events that are truncated by the camera border are reconstructed ***incorrectly*** with an ALPHA close to 90°. Consequently, for this method to work properly, the distribution must be ***forced*** to become as flat as possible. This can be accomplished , for example, by excluding events with a large DIST.

A.3.5 The linear discriminant analysis as a dynamical cut and as a tool to quantify the discrimination power of image parameters

The ***linear discriminant analysis*** (***LDA***) is used to find optimized and unbiased linear combinations of variables in order to separate classes of events [Fab97]. The method was invented by Fisher [Fis37]. In this case, the classes are signal and background events. It deals naturally with dependencies between variables and is not useful if they are NOT dependent on each other. In this case, normal cuts give better results. In this section I show how a combination of static cuts and the LDA achieves a good separation.

As practical side information, an estimator, called ***discriminating power***, of how well two samples can be separated, is obtained. The linear discriminant method is nothing more than a simple ***linear*** neural net with only ***one neuron***, a so-called perceptron with ***linear output***, trained by an analytical method (linear least squares method), rather than doing stepwise optimization.

The main ***advantage*** of the LDA, compared to a neural net and to other cut-optimization methods as well, is that it is ***analytically optimized*** while all other methods need an ***iterative*** maximization procedure whose result depends on ***initial values*** and which are often ***not reproducible***.

It is exactly this reproducibility that is needed when seeking an ***estimator*** that quantifies how good a given parameter set is able to separate two classes of events.

Description of the method

The method finds ***discriminant factors*** which are ***linear combinations*** of the input parameters. For the case of ***two*** classes, there is only one discriminant factor which is called ***discriminant variable*** and which is the optimum linear combination for the separation between the two classes. The method is first described for the case of q classes and then simplified to two classes.

Each class has N_i events. The total number of events is:

$$N_{tot} = \sum_i N_i \tag{A.43}$$

Let $\mathbf{x}^i(n) = \left(x_1^i(n), \ldots, x_p^i(n)\right)$ be a vector of p input parameters which represents the event n in class i. The ***expectation value*** of parameter x_j in class i is

$$E\left[x_j^i\right] = E_j^i = \frac{1}{N_i}\sum_{n=1}^{N_i} x_j^i(n) \tag{A.44}$$

and for all events of ***all classes combined*** we have:

$$\begin{aligned} E\left[x_j\right] = E_j &= \frac{1}{N_{tot}}\sum_{i=1}^{q}\sum_{n=1}^{N_i} x_j^i(n) \\ &= \frac{1}{N_{tot}}\sum_{i=1}^{q} N_i E_j^i \end{aligned} \tag{A.45}$$

The overall ***covariance matrix*** element between the variables x_k and x_l for all classes combined is given by

$$T_{kl} \equiv covar\left[x_k, x_l\right] = E\left[\left(x_k - E\left[x_k\right]\right) \times \left(x_l - E\left[x_l\right]\right)\right] \tag{A.46}$$

The covariance matrix can be ***split into two parts***, the '***within' matrix*** $\mathbf{W}$ and the '***between' matrix*** $\mathbf{B}$

$$T_{kl} = W_{kl} + B_{kl} \tag{A.47}$$

whereas

$$W_{kl} = \frac{1}{N}\sum_{i=1}^{q}\sum_{n=1}^{N_i}\left(x_k^i(n) - E\left[x_k^i\right]\right) \times \left(x_l^i(n) - E\left[x_l^i\right]\right) \tag{A.48}$$

$$B_{kl} = \sum_{i=1}^{q}\frac{N_i}{N_{tot}}\left(E_k^i - E_k\right) \times \left(E_l^i - E_l\right) \tag{A.49}$$

$\mathbf{W}$ describes the variances of one parameter within the group whereas B describes the differences between groups. Let us now introduce the linear combination of the input parameters

$$u(n) = \sum_{j=1}^{q} y^j x_j(n) = \mathbf{y}^T\mathbf{x}(n) \tag{A.50}$$

where $\mathbf{y}^T$ is the transposed matrix of $\mathbf{y}$ and $\mathbf{y}$ is the equivalent of ***weight vectors*** in the case of neural nets. The variance of $u(n)$ is given by

$$var(u) = \frac{1}{N_{tot}}\sum_{i=1}^{q}\sum_{n=1}^{N_i}\left(u^i(n) - E\left[u^i\right]\right) \tag{A.51}$$

$$= \mathbf{y}^T\mathbf{T}\mathbf{y} \tag{A.52}$$

$$= \mathbf{y}^T\mathbf{W}\mathbf{y} + \mathbf{y}^T\mathbf{B}\mathbf{y} \tag{A.53}$$

The first term in the last line is equivalent to the variance ***within*** the classes while the second term describes the differences ***between*** the classes. To optimize the discrimination, the first term must be ***minimized*** while the second one must be ***maximized***. The discriminating power is defined as

$$f(u) \equiv \frac{\mathbf{y}^T\mathbf{B}\mathbf{y}}{\mathbf{y}^T\mathbf{T}\mathbf{y}} \tag{A.54}$$

and the ***optimization*** is done by deriving $f(u)$ and setting it equal to zero:

$$\frac{\partial f(u)}{\partial(y_j)} = 0$$

This leads to the expression

$$\mathbf{T}^{-1}\mathbf{B}\mathbf{y} = f(u)\mathbf{y} \tag{A.55}$$

which means that $f(u)$ is maximal if it is an ***eigenvalue*** of $\mathbf{T}^{-1}\mathbf{B}$ and that $\mathbf{y}$ (the weight vector) is the corresponding eigenvector.

Now in our case of ***two classes***: We have gamma events and hadronic events. They will be denoted with indices G and H, respectively. The ***expectation values*** of the two classes for each parameter can be expressed as a vector:

$$\mathbf{E} = \left(E_1^i, ..., E_p\right)^T \tag{A.56}$$

The matrix becomes

$$\mathbf{B} = \mathbf{v}^T\mathbf{v} \tag{A.57}$$

	ALPHA	WIDTH	LENGTH	CONC	DIST	log SIZE	$\cos\theta$	RMS
ALPHA	100	-0.5	-0.1	1.0	0.4	-0.5	-0.3	-0.1
WIDTH		100	42.6	-59.6	-5.9	36.5	15.4	-7.6
LENGTH			100	-70.0	-21.1	35.7	10.3	-3.8
CONC				100	22.0	-47.7	-14.7	5.8
DIST					100	-7.5	0.3	0.5
log SIZE						100	5.6	-1.6
$\cos\theta$							100	-20.0
RMS								100

Table A.1: The table shows the correlation in percent for the most important image parameters. WIDTH, CONC and LENGTH are strongly correlated. SIZE and CONC also show a dependence. Interesting is the correlation between zenith angle and NSB (pedestal RMS). ALPHA is not correlated with any of the image parameters.

with

$$\mathbf{v} = \frac{\sqrt{N_G N_H}}{N_{tot}} \left(\mathbf{E}^G - \mathbf{E}^H\right) \tag{A.58}$$

Now the solution to the maximization problem (Equ. A.55) for our ***weights*** $\mathbf{y}$ gets the simple form

$$\mathbf{y} = \mathbf{T}^{-1}\mathbf{v} \tag{A.59}$$

The ***discrimination power*** is a very useful quantity as we will see later. It can be used to quantify how well the datasets have been discriminated. It shows how well the parameter-set is able to discriminate between the two classes. Its value ranges from 'zero' to 'one'. 'One' means perfect separation and 'zero' means no separation.

$$D = f(u) = \mathbf{v}^T\mathbf{y}\,;\ D \in [0,1] \tag{A.60}$$

Equ. A.50 can be used to calculate the center (offset of origin) between the two distributions $u^i(n)$ for gammas and hadrons (i=G,H) by taking the ***average of the two expectation vectors***:

$$u_0 = \frac{1}{2}\mathbf{y}^T\left(\mathbf{E}^G + \mathbf{E}^H\right) \tag{A.61}$$

Finally, the ***output of the LDA*** (similarly to the output of a neural net) takes the form

$$LDA_{out}(n) = u(n) - u_0 = \mathbf{y}^T\mathbf{x}(n) \tag{A.62}$$

Before applying the method, the ***image parameters*** that enter as inputs to the method must be ***chosen.***

Application of the LDA to Monte Carlo Data samples

As a first step it must be stated that the separation is ***best*** if the number of events in both samples are ***equa***l. If this is not the case, the datasets have to be ***normalized*** to the same number of events.

As explained above, the LDA is expected to achieve an ***improvement in separation*** only in the case of ***correlated*** variables. It uncorrelates dependencies. All parameters that are ***not*** correlated should be treated by a ***static cut***. Tab. A.1 shows the ***correlation matrix*** for a real data sample (mainly hadronic events).

The correlation matrix for MC gamma events looks similar but is not exactly the same. The table shows that ***ALPHA*** has absolutely ***no correlation*** to the other parameters at all. It is used separately to ***estimate*** the background. It is interesting to see that there is ***significant correlation*** between the night sky background (***pedestal RMS***) and the ***zenith***

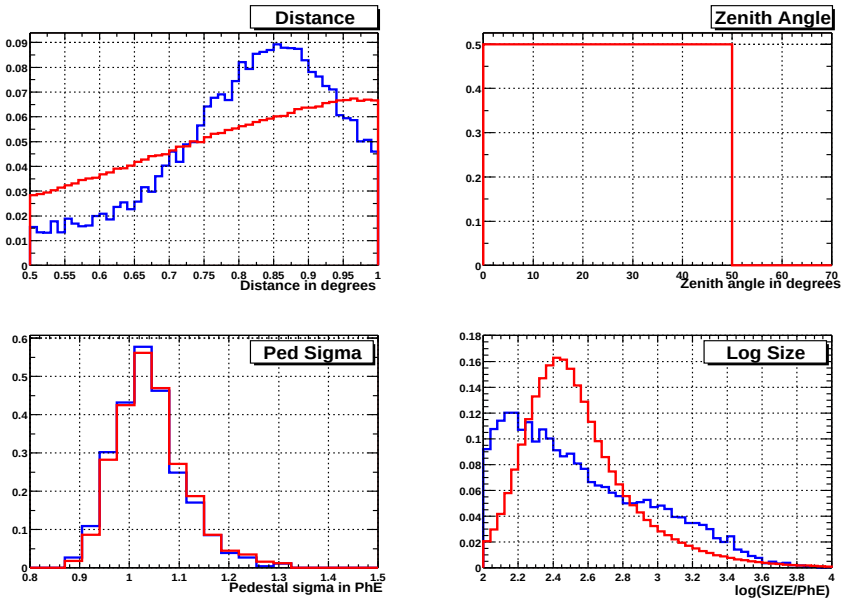


Figure A.9: Distributions of the four parameters SIZE, DIST, Pedestal RMS and the zenith angle, for gamma events (blue) and hadronic events (red) after normalization. The SIZE distribution shape for hadrons and gammas is quite different because of a) a different spectral index of the γ-MC and the OFF-data and because of b) different effective areas for hadrons and γ's. The pedestal RMS (NSB) distribution is almost the same because the NSB has been simulated to the γ-MC in a way that it is equal to the OFF-dataset.

angle. The correlation between ***SIZE*** and the other three parameters: ***WIDTH, LENGTH*** and ***CONC*** is logical. The parameter ***CONC*** has such a high degree of correlation that it is essentially represented by ***WIDTH*** and ***LENGTH*** .

The image parameters can be divided into two groups:

1. ***cut parameters*** for background suppression and
2. ***non-cut parameters*** e.g. estimators.

The distributions of the ***non-cut parameters*** as ***SIZE***, ***DIST***, ***pedestal RMS*** and the ***zenith angle*** have to be ***equal*** for both, gamma events and hadronic events, because we ***do not*** wish to cut on them. Still they are included in the LDA input parameters to ***correct for dependencies*** of the cut parameters on them (similar to dynamic cuts).

To make the distributions of the non-cut parameters equal

1. the night sky background is simulated in the MC gamma sample in a manner such that it has the ***same distribution*** with respect to ***zenith angle*** as the OFF-data set (which is used as hadronic events).
2. The z***enith angle distribution*** was normalized to be ***flat*** in both data sets since it is important to have equal cut efficiencies for all zenith angles.

In this analysis the ***SIZE*** distributions and the ***DIST*** distributions are ***not normalized*** to each other. The experience showed that the separation does not improve when they are normalized. Fig. A.9 shows the distribution of these four parameters after normalization.

If the parameters are included ***linearly***, the LDA represents a ***cut plane*** in a ***multi-dimensional space***, separating the two datasets (represented by vectors) maximally with

respect to discriminating power. Figure A.10 shows the output of equ. A.62 after optimization (equ. A.59and equ. A.60). The discriminating power is 0.62. The parameters included are:

$$\begin{aligned} Input[0] &= WIDTH \\ Input[1] &= LENGTH \\ Input[2] &= CONC \\ Input[3] &= SIZE \\ Input[4] &= \cos\theta \end{aligned} \tag{A.63}$$

Again, ***ALPHA*** is not included in this list, since it is a ***independent*** parameter. It is used to ***estimate*** the background.

A cut in the ***one dimensional output*** of the LDA can be chosen. Depending on the position of the cut, a higher or lower percentage of background is removed while retaining the signal events respectively. Fig. A.11 shows the cut efficiencies ε_G and ε_H against each other. Fig. A.12 demonstrates the performance of the LDA for a cut which achieves maximal significance. The cut efficiency increases for increasing energy and remains ***flat*** for increasing zenith angle, as it should be.

If ***higher orders and combinations*** of image parameters are included in the LDA then it no longer corresponds to a linear cut plane, but rather to a ***multidimensional surface*** that follows the multi-dimensional ***shape*** of the parameter distributions in a more ***adaptive*** manner (in a similar way as dynamic cuts). Consequently, the separation improves. The following terms have been appended to the previous parameter list:

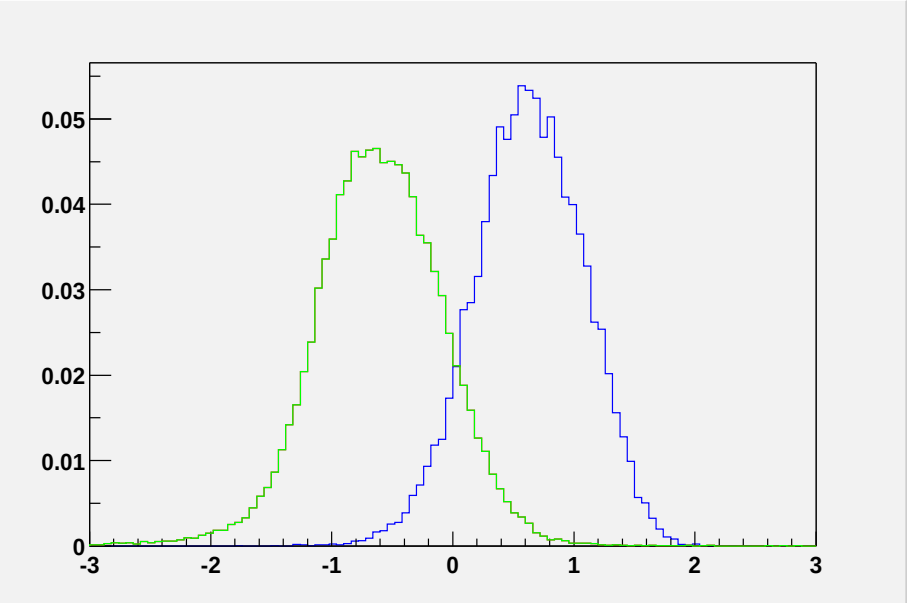
$$\begin{aligned} Input[5] &= WIDTH^2 \\ Input[6] &= LENGTH^2 \\ Input[7] &= CONC^2 \\ Input[8] &= \frac{WIDTH * LENGTH}{SIZE} \end{aligned} \tag{A.64}$$

Fig. A.13, Fig. A.11 and Fig. A.10 show the improvement in separation using higher orders and a combination of these parameters. The excess events found with linear input gives 11400 +- 190. This increases to 12600 +- 200 for the LDA with higher order inputs.

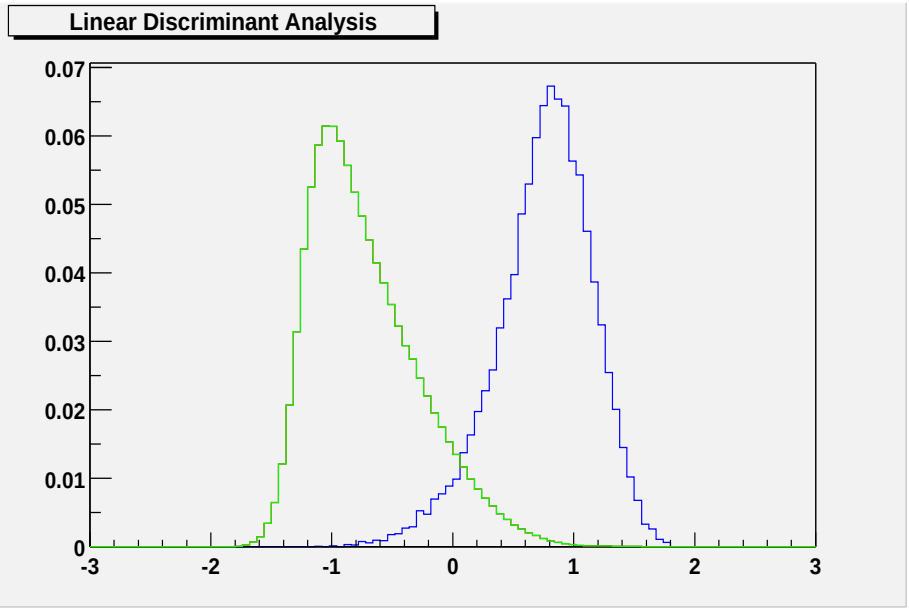
The Fig. A.14 shows the excess events found as a function of the cut position of the LDA output. A ***plateau*** appears on the left side of the plot which can be interpreted as the ***total signal events*** in this data set. A fit on the plateau gives $N_{Ex} = 15030 \pm 71$. When compared with the number of excess events found at maximum significance (12600 +- 200), this number results in a ***total cut efficiency of 83 %***. It must be noted that before application of the selection cut, a ***'precut'*** on the minimum SIZE < 60 PhE, on the zenith angle < 50° and on the distance 0.4 < DIST < 1.05 was performed, thus previously reducing the total number of signal events in the dataset. It is clear that this plot here does ***not*** tell us anything about the ***trigger efficiency*** nor about the efficiency of the 'precut'.

A.3.6 Comparison of the discrimination power of the LDA and the dynamical cuts, tested with real Mkn 421 data

For the comparison of ***dynamical cuts*** of the analysis program 'jacuzzi' at MPI Munich [Kra01] with the ***LDA*** of this thesis, the ***same*** dataset was used ***without*** pointing corrections (pointing corrections increase the significance of the signal). For case of the dynamical cuts a correction of the parameter with respect to the NSB (the so-called 'zonking' [Kes01]) and also a cut on the ***asymmetry*** was applied (see the following section). Both

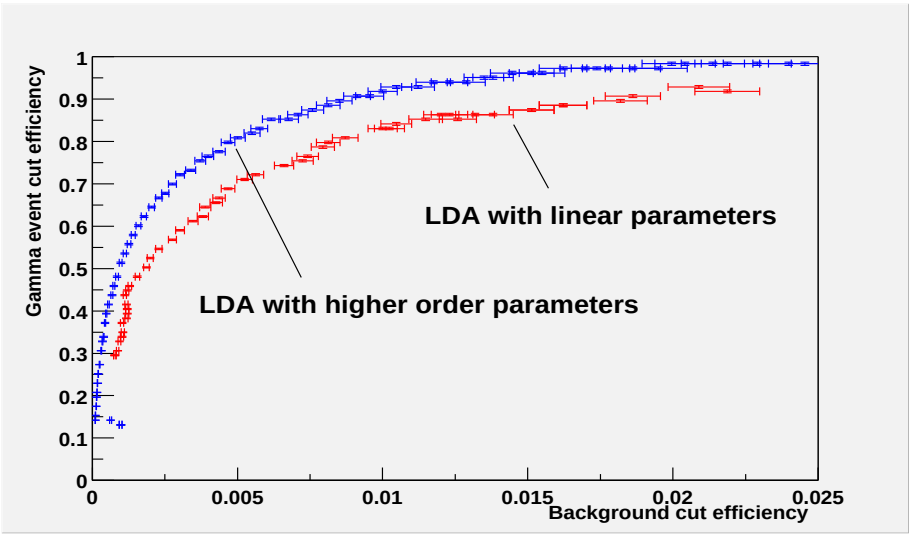


(a) Linear inputs

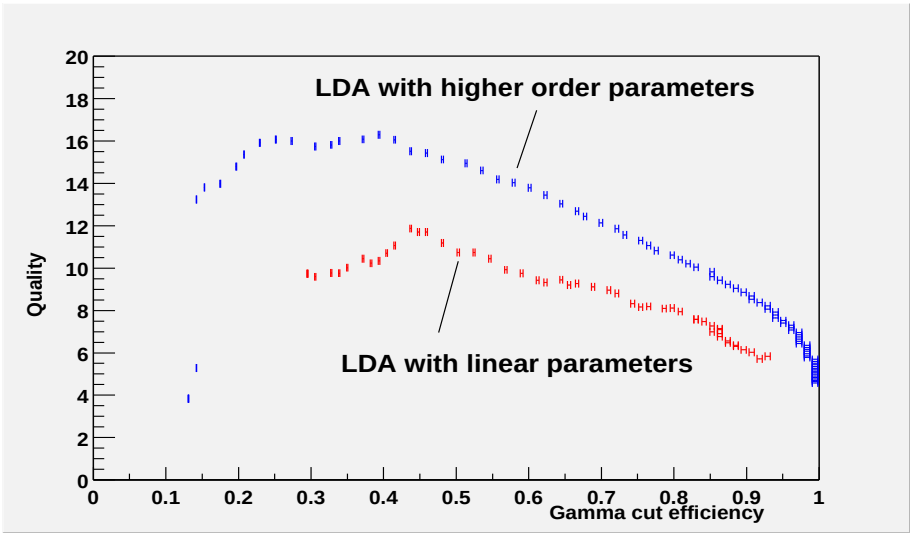


(b) Higher order inputs

Figure A.10: Output of the LDA A.62 for gamma events (blue) and hadronic events (green). The areas are normalized to one. In (a) the parameters are included linearly (D=0.62) and in (b) higher polynomial orders are included as well (D=0.748). It is clearly shown that the discrimination increases by including higher orders of the parameters.



(a)



(b)

Figure A.11: (a) Gamma cut efficiency vs background cut efficiency and (b) Quality factor for the LDA vs LDA-output. Both plots show the performance of LDA with only linear inputs and with higher order inputs. The inclusion of higher orders in the LDA improve the quality factor and its separation capability significantly.

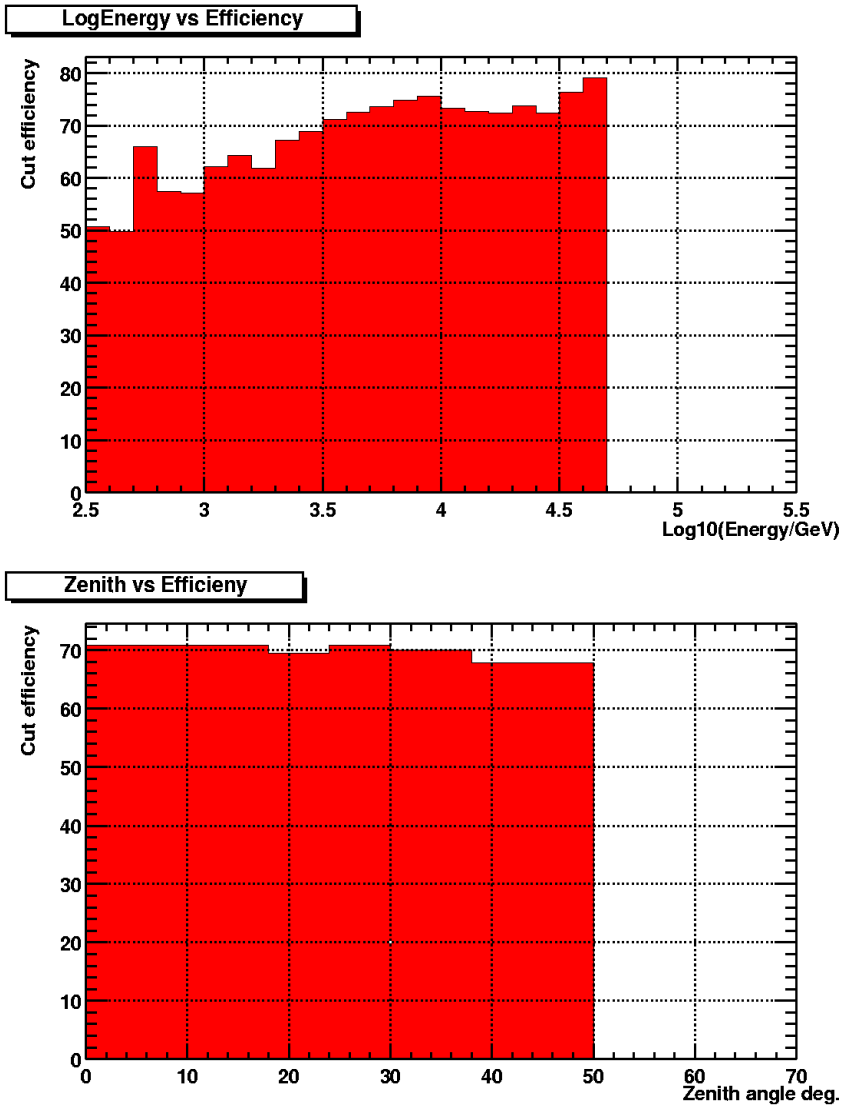
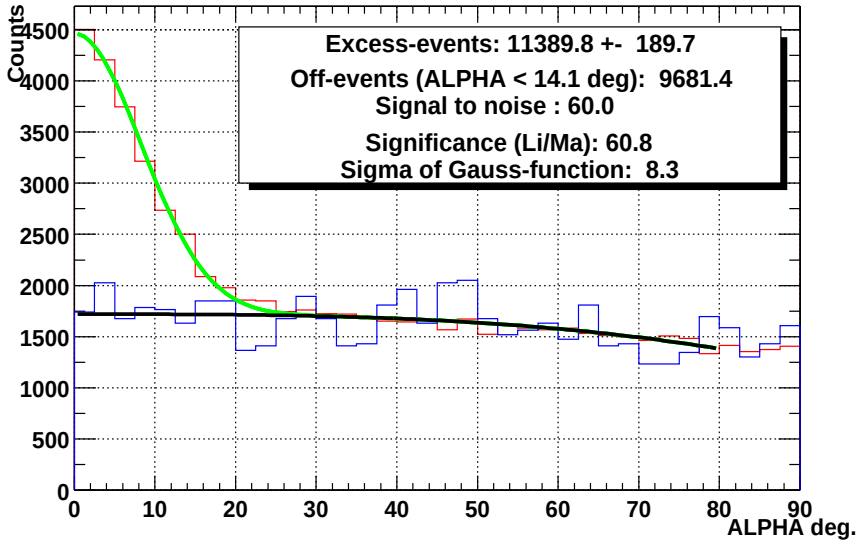
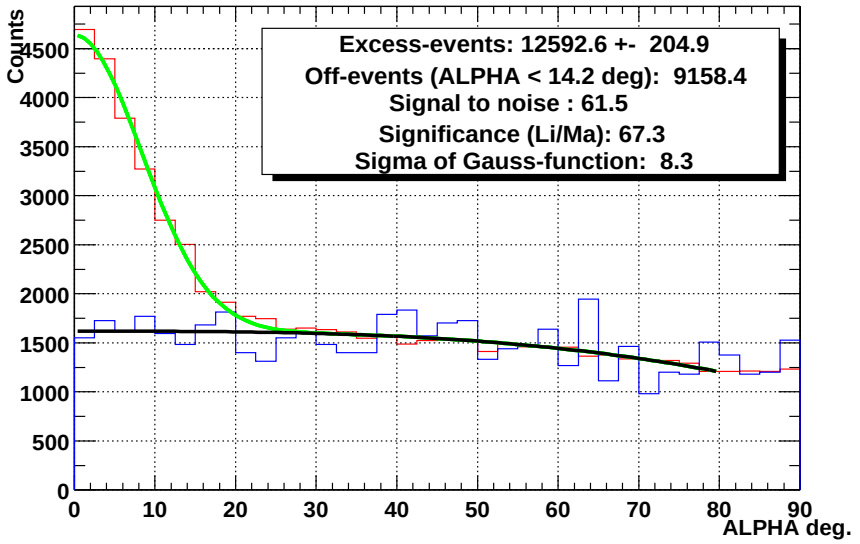


Figure A.12: These two plots show the γ cut efficiency for the LDA with parameter input of only linear order and for a cut which achieves maximal significance. In the upper the γ cut efficiency is plotted against the logarithm of the energy in log(GeV) and in the lower plot against the zenith angle. The cut-efficiency increases with energy and is flat with zenith angle as it should be.



(a)



(b)

Figure A.13: The linear discriminant analysis is applied to a sample Mkn 421 dataset of 167 h observation time. a) The upper plots shows the result for the LDA with linear inputs. It yields in a significance (Li/Ma) of 60.8. b) The lower plots shows the result for the LDA with higher orders of image parameters for the input. The background is more suppressed while more signal is found (increase from 11400 events to 12600 events). The significance improves by 10 %. Here a previous filter cut on SIZE > 60 PhE, the 0.4 DIST > 1.05 and on the zenith angle <50° was applied.

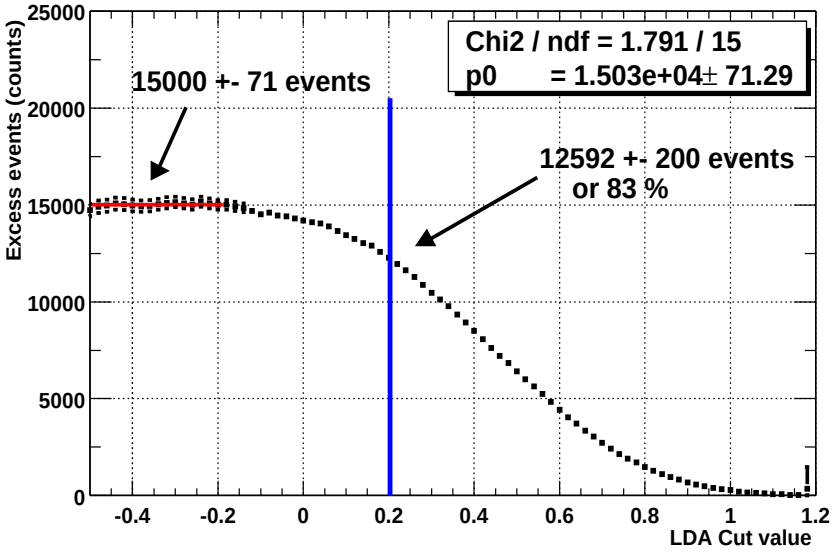


Figure A.14: When moving the LDA cut position a plateau appears on the left side which indicates the maximum excess events in this dataset. This maximum represents the total signal events in this dataset. Here a previous cut on SIZE > 60 PhE, the 0.4 DIST > 1.05 and on the zenith angle <50° was applied. After the LDA selection (higher order inputs) about 83 % of the total events have been selected.

improvements ***have not been*** included in the ***LDA*** procedure up to now. The result of the dynamical supercuts can be seen in Fig. A.15.

As a result of not applying pointing corrections, the ALPHA distribution is wider and unfortunately, 'jacuzzi' was fitting the ALPHA distribution with a fixed variance for the Gaussian distribution, expecting a narrower distribution (it was not made for that). Consequently, the fit has large chisquare. Still, the fit result can be ***roughly*** compared with the result achieved by the LDA. The excess events found by 'jacuzzi' are 7700 +- 130 and the LDA (in this simple version) found 12600 +- 200. The significance of the LDA is around 20% higher. It has to be kept in mind that jacuzzi also finds higher significances when a pointing correction has been applied beforehand.

As a ***conclusion*** it can be stated that the LDA is at least as efficient than dynamical cuts.

A.3.7 Conclusion about the efficiency of the LDA separation method and an outlook

The LDA proves to be a very efficient separation method that can well compete with classic methods like static cuts and the dynamic cuts and has ***equal*** or better ***performance***. It shows to be especially efficient for very low energy events (< 1 TeV) and very high energy events (> 10 TeV).

The ***main advantage*** of the ***LDA*** is clearly the cut optimization algorithm that is ***analytically solvable***. This does not only speed up the optimization process but it does also ***not depend*** on any initial value as any other stepwise optimizing methods (e.g. 'MINUIT'-package). Its results are 100 % reproducible unlike any other kind of separation methods as dynamic cuts or neural networks.

For this reason it is especially suited to ***quantify the separation power*** of different image parameters sets in order to find the most efficient one. The LDA will be used in the

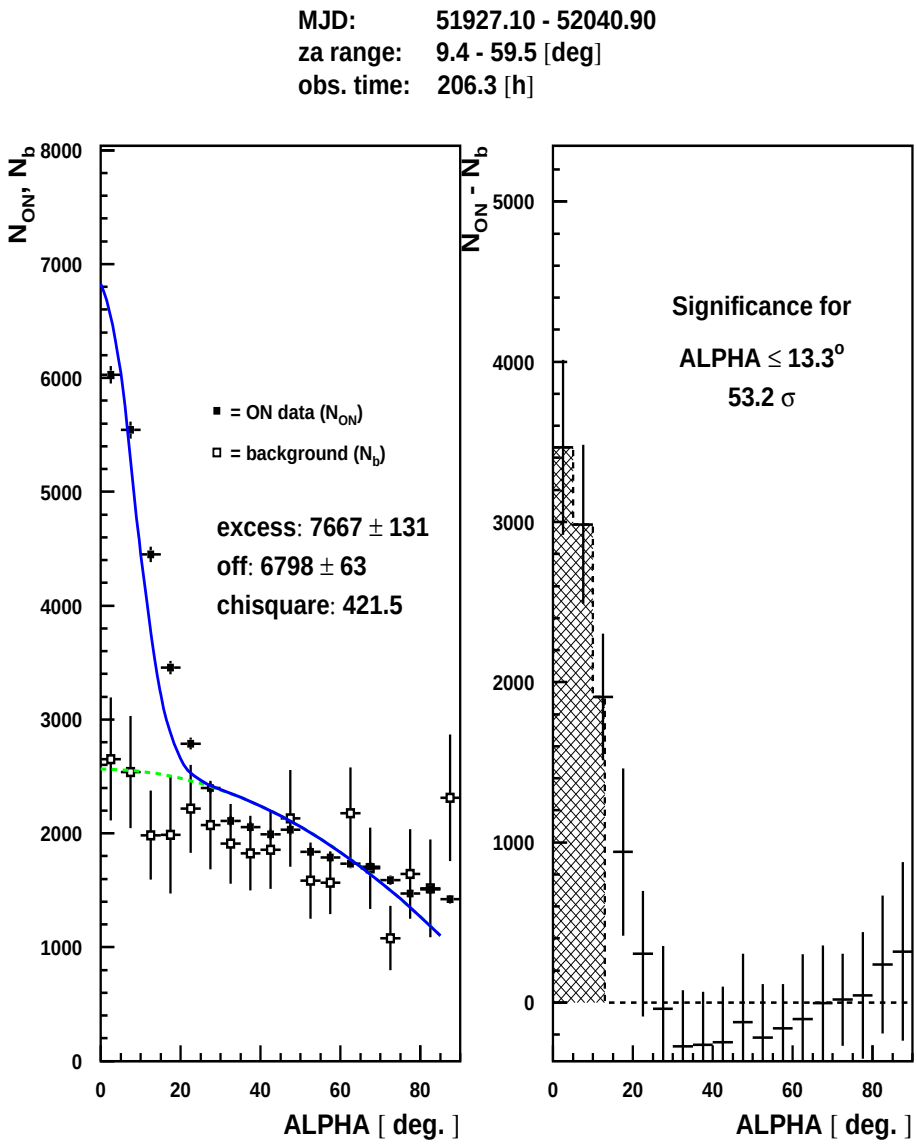


Figure A.15: Significance determined by the 'jacuzzi' program (dynamic selection cuts) as applied on the whole Mkn 421 dataset from February 2001 to April 2001 without pointing correction.

following section for this purpose.

For ***future improvements*** of the performance the author suggests to interested people to ***introduce a binning in energy*** and optimize the LDA in ***each*** energy bin separately. This will solve the difficulty of having ***different*** energy spectra for the ***background*** and the ***signal*** and will further ***increase*** the separation power.

A.4 New image parameters and cleaning algorithms

The LDA has been defined and can be used to check the separation ability of ***new*** image parameters. Each new parameter can be included as an ***additional parameter*** in the LDA.

There are ***two*** estimators that tell us about the ***separation power*** of the parameters. These are the ***discriminant power*** (Equ. A.60) and the ***final significance*** can be compared with different parameter sets, upon application to the data. With both we can ***judge*** whether an improvement has been achieved. Three things were tested here.

1. ***New*** image parameters were developed which describe more features in the image structure,
2. ***Weights*** were introduced in the calculation of Hillas parameters and
3. The ***cleaning level*** of the 'dynamical' image cleaning has been varied systematically in order to find the one which results in the highest separation power for the image parameters.

A.4.1 Introduction of weights in the calculation of Hillas parameters

Up to now, in the definition of the Hillas parameters presented above, Equ. A.25 to Equ. A.33the weight has always been assumed to be equal to the charge in one pixel ($w_i = q_i$). Here weights of the following form have been tested:

$$w_i = q_i^n \quad (NONE) \tag{A.65}$$

$$w_i = \frac{q_i^{n+0.5}}{\sqrt{q_i + \sigma_i^2}} \quad (REL) \tag{A.66}$$

In Equ. A.65 (these weights are denoted in the following as '***NONE***'), q_i is potentiate with an exponent n. This has the effect that pixels with ***more signal*** have ***more weight*** in the calculation of the HILLAS variances. This makes them less sensitive to noise fluctuations (coming from NSB).

Equ. A.66 (these weights are denoted as '***REL***') a weight is introduced that represents the ***statistical error*** of the signal in the pixel ($\sigma_{tot,i} = \sqrt{q_i + \sigma_i^2}$). A value 0.5 has been added to the exponent to keep the ***effective exponent*** in both cases ***equal***. The exponent become comparable in this manner. For a small NSB the weights 'REL' converge to the weights 'NONE'.

A program has been written which ***systematically compares*** the separation power of a MC Gamma dataset and a hadronic background dataset that was recorded by CT1. The ***NSB*** has been simulated in the MC gamma dataset to have a zenith angle distribution ***equivalent*** to that of the hadronic dataset (See section A.2.1).

The program tries systematically

1. Different ***image cleaning levels***,
2. Different ***cleaning types*** and
3. Four different ***exponents*** (n=0.5, 1.0, 1.5, 2.0).

This is done for ***both*** types of weights , 'NONE' and 'REL' (Equ. A.65and Equ. A.66).

To be able to judge the separation, ***two*** parameters were introduced. I called them:

1. ***Separation***

$$S = \frac{\langle p_G \rangle - \langle p_H \rangle}{\sqrt{\sigma_G^2 + \sigma_H^2}} \tag{A.67}$$

   which is the difference in the mean value of the parameter divided by its combined variance. It should be as ***large*** as possible for good separation. This parameter does not take into account the ***shape*** of the distributions.

2. The ***overlapping integral*** is ***sensitive*** to the shape of the distributions

$$\begin{aligned} O &= \frac{\int p_G(x)\, p_H(x)\, dx}{\int p_G(x)\, dx \int p_H(x)\, dx} \\ &= \frac{\sum_i \left(p_G^i\, p_H^i\right)}{\left(\sum_i p_G^i\right)\, \left(\sum_i p_H^i\right)} \end{aligned} \tag{A.68}$$

   It measures the ***overlap*** of the two distributions and should be as ***small*** as possible for good separation.

With the combination of these two parameters we have a ***tool*** to judge possible improvements in ***separation*** for different configurations.

The results for the parameter WIDTH, LENGTH, CONC are shown in Fig. A.17, Fig. A.16, Fig. A.18, Fig. A.19, and A.20. The exponent is expressed in different colors. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles*** and a ***continuous line*** represent ***'classic'*** cleaning, ***triangles*** and a ***dashed line*** are '***island'*** cleaning and ***stars*** with ***dotted line*** are ***'mountain'*** cleaning.

Separation of WIDTH, LENGTH and CONC

Tab. A.2 illustrates the ***cleaning levels*** for the ***best*** separations and overlaps for the ***two*** different cleaning algorithms, ***'classical'*** and ***'island'*** cleaning. The weights REL and NONE show almost ***identical*** results (see Fig. A.17, Fig. A.16, Fig. A.18 and Fig. A.19). The first two rows show the old classic version of HILLAS parameters. The cleaning level consists of two numbers, the 'image core limit' and the 'image border limit'. The best cleaning algorithms for WIDTH, LENGTH and CONC are printed in bold. A cut of $0.4 < DIST < 1.05$, $SIZE > 100$ and $12\,deg \leq Zenith\,angle \leq 22\,deg$ was applied beforehand ('precut').

As a conclusion can be stated the following:

1. There is ***no visible difference*** between the weights 'NONE' and 'REL'

2. ***Slightly reducing*** the cleaning level to 2.7/2.0 ***increases*** the separation a little for ***'classic'*** cleaning.

3. The ***best separation*** and ***overlap*** (for all three parameters) is achieved for the ***'island' cleaning method*** with a ***very low*** cleaning level of 1.0/0.3.

4. In case of the parameter ***WIDTH*** the separation ***increases*** by increasing the ***exponent*** n from $n = 1$ to $n = 1.5$.

5. In case of ***LENGTH*** $n = 1$ remains the best choice.

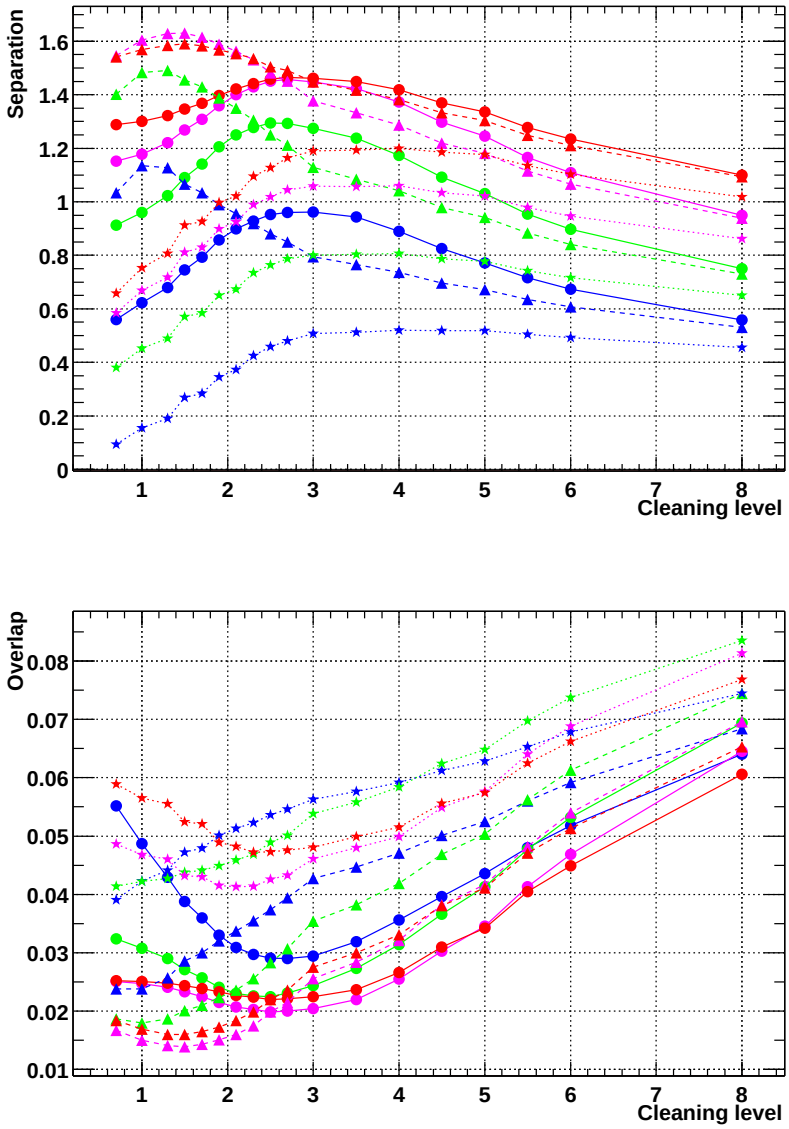


Figure A.16: This figure shows the ***separation*** and the ***overlap*** for the parameter ***WIDTH*** for ***NONE***-weights (Equ. A.65) for different 'core' cleaning levels. The ***exponent*** n is coded in ***colors***. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles and continuous line*** represent ***'classic'*** cleaning, ***triangles and dashed line*** is ***'island'*** cleaning and ***stars with dotted line*** are '***mountain'*** cleaning. The best separation is obtained with 'island' cleaning, an exponent of 1.5 and a cleaning level of 1.5.

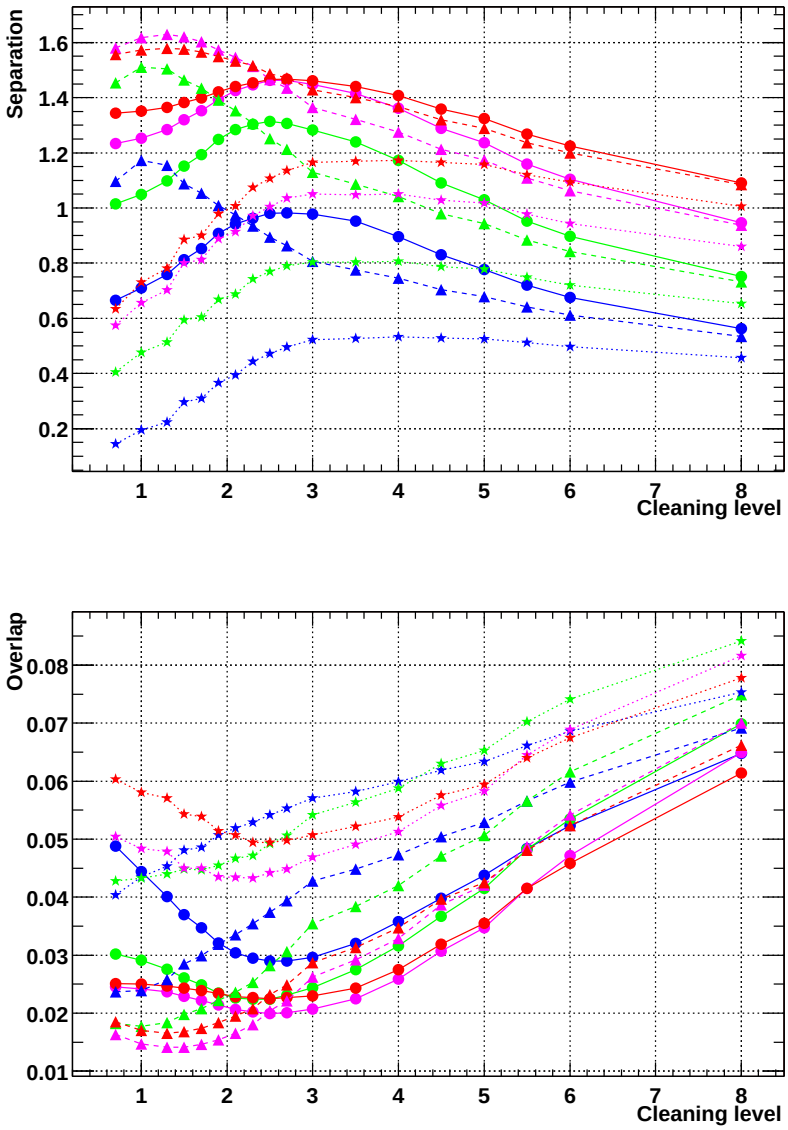


Figure A.17: This figure shows the ***separation*** and the ***overlap*** for the parameter ***WIDTH*** for ***REL***-weights (Equ. A.66) for different 'core' cleaning levels. The ***exponent*** n is coded in ***colors***. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles and continuous line*** represent ***'classic'*** cleaning, ***triangles and dashed line*** is ***'island'*** cleaning and ***stars with dotted line*** refer to '***mountain'*** cleaning. The best separation is obtained with 'island' cleaning, an exponent of 1.5 and a cleaning level of 1.5. REL weights give results identical to NONE weights.

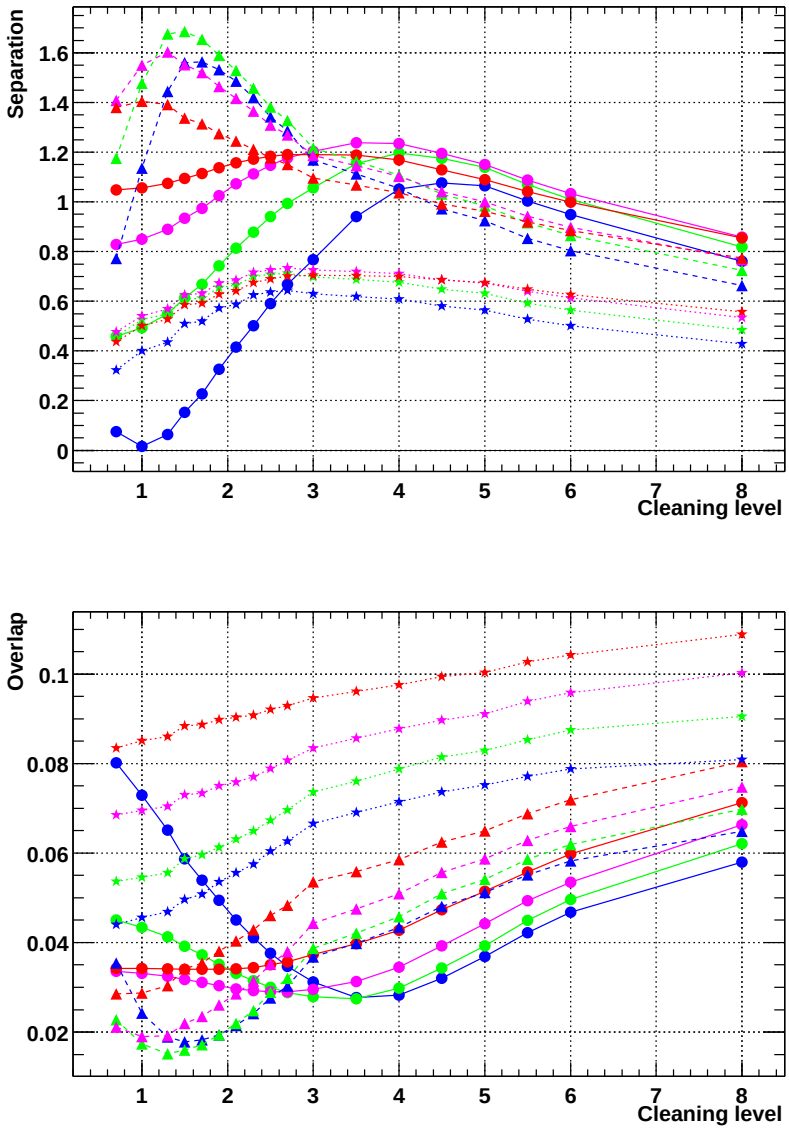


Figure A.18: This figure shows the ***separation*** and the ***overlap*** for the parameter ***LENGTH*** for ***NONE***-weights (Equ. A.65) for different 'core' cleaning levels. The ***exponent*** n is coded in ***colors***. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles and continuous line*** represent ***'classic'*** cleaning, ***triangles and dashed line*** refer to ***'island'*** cleaning and ***stars with dotted line*** are '***mountain'*** cleaning. The best separation is obtained with 'island' cleaning, an exponent of 1.0 and a cleaning level of 1.5.

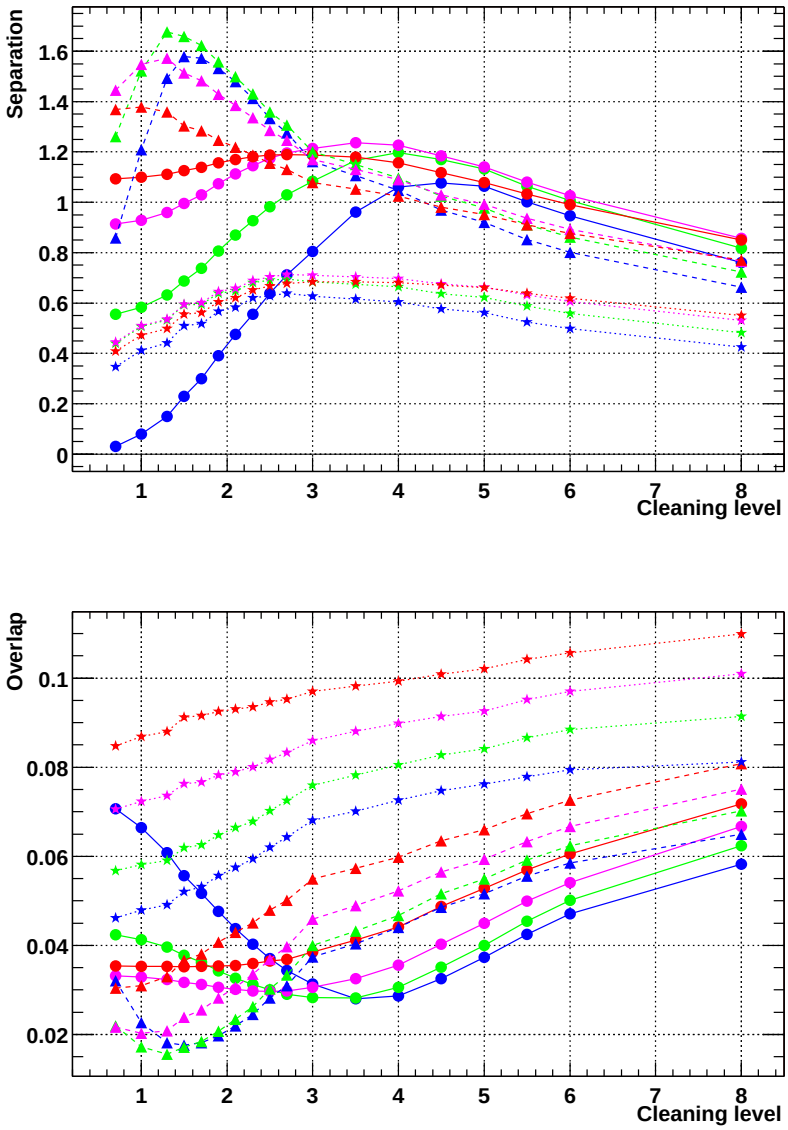


Figure A.19: This figure shows the ***separation*** and the ***overlap*** for the parameter ***LENGTH*** for ***REL***-weights (Equ. A.66) for different 'core' cleaning levels. The ***exponent*** n is coded in ***colors***. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles and continuous line*** represent ***'classic'*** cleaning, ***triangles and dashed line*** is ***'island'*** cleaning and ***stars with dotted line*** are '***mountain'*** cleaning. The best separation is obtained with 'island' cleaning, an exponent of 1.0 and a cleaning level of 1.5. REL-Weights give results identical to NONE-weights.

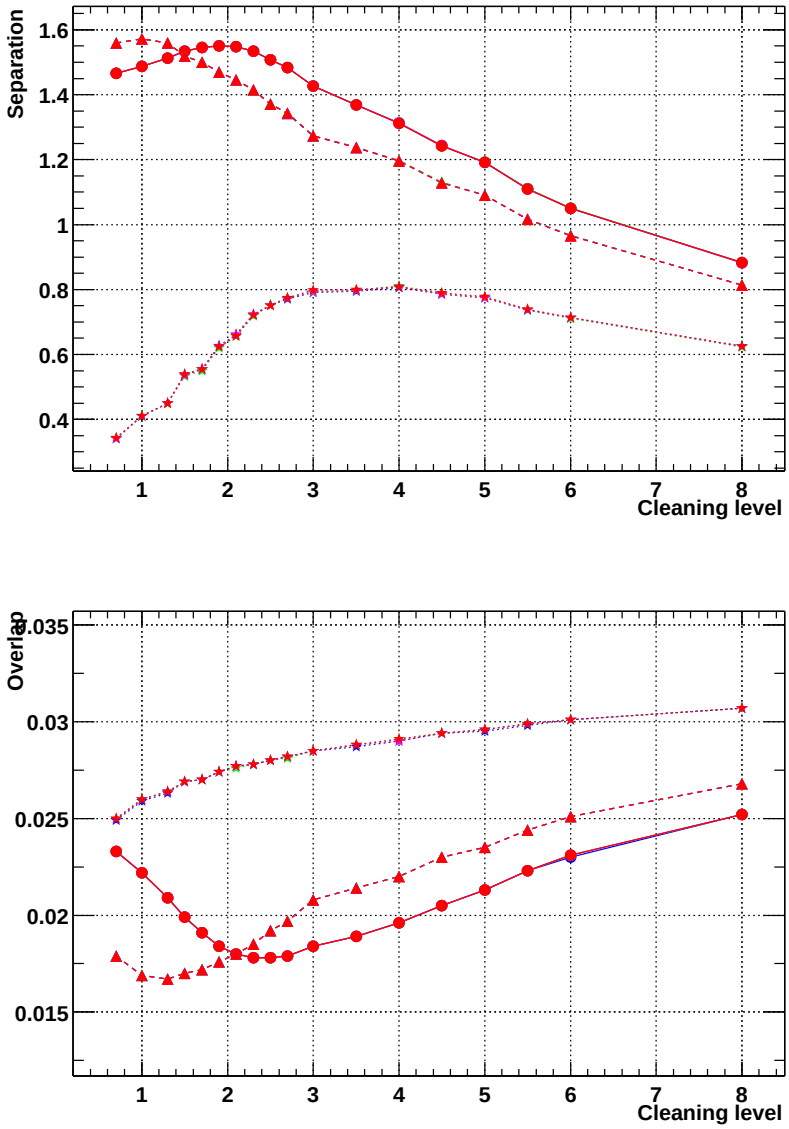


Figure A.20: This figure shows the ***separation*** and the ***overlap*** for the parameter ***LENGTH*** for ***NONE***-weights (Equ. A.66) for different 'core' cleaning levels. The ***exponent*** n is coded in ***colors***. Different cleaning algorithms are expressed in different marker styles and line styles. ***Circles and continuous line*** represent ***'classic'*** cleaning, ***triangles and dashed line*** is ***'island'*** cleaning and ***stars with dotted line*** are '***mountain'*** cleaning. The best separation is obtained with 'island' cleaning and a cleaning level of 1.0.

Parameter	Cleaning type	Level	Exp.	Sep.	Overlap
WIDTH	'classic'	3.0/2.5	1.0	1.278	0.0238
LENGTH	'classic'	3.0/2.5	1.0	1.057	0.0279
WIDTH	'classic'	2.7/2.0	2.0	1.464	0.0200
WIDTH	**'island'**	**1.3/0.3**	**1.5**	**1.630**	**0.0139**
LENGTH	'classic'	3.5/2.7	1.5	1.235	0.0345
LENGTH	**'island'**	**1.3/0.3**	**1.0**	**1.676**	**0.0151**
CONC	***'island'***	***1.0/0.3***	/	***1.584***	***0.0180***

Table A.2: The table shows the ***cleaning levels*** for the ***best*** separations and overlaps for WIDTH, LENGTH and CONC for the ***two*** cleaning algorithms, ***'classical'*** and ***'island'*** cleaning. The weights REL and NONE show almost ***identical*** results (see Fig. A.17, Fig. A.16, Fig. A.18 and Fig. A.19). The first two rows show the old 'classic' version of HILLAS parameters. The cleaning ***level*** consists of two numbers, the 'image core limit' and the 'image border limit'. 'Exp' denotes the exponent n and 'Sep' is the separation.
The best cleaning algorithms for WIDTH and LENGTH are printed in bold.

Cleaning type	Level	Exponent	Mean	Variance
'classic'	3.0/2.5	1.0	12.8	11.8
'classic'	3.0/2.5	1.5	10.0	8.4
'classic'	4.0/3.0	1.5	8.4	6.4
'island'	***4.0/3.0***	***1.5***	***3.9***	***3.9***

Table A.3: The table shows the best conditions to obtain the sharpest ALPHA distribution for point sources. The mean and the variance are the quantities that quantify the sharpness. The first row shows the sharpness of the 'classic' ALPHA as defined by HILLAS. The sharpest ALPHA distribution is obtained by increasing n to $n = 1.5$ and by removing the 'islands' with the 'island' cleaning algorithm.

Conditions to obtain the sharpest ALPHA distribution

The same procedure was applied to analyze the ***conditions*** to obtain the ***sharpest*** ALPHA distribution (for γ-events and for point sources). The two ***quantities*** that characterize the ***sharpness*** are the ***variance*** and the ***mean*** of the ALPHA distribution and they should be kept as ***small*** as possible. The ***smaller*** these two the quantities, the ***better*** the background suppression. Fig. A.21 illustrates both values for different ***cleaning levels***, cleaning ***algorithms*** and different ***exponents*** n, as before. The colors and markers are coded in the same way as described above.

A cut of $0.4 < DIST < 1.05$, $SIZE > 100$ and $12\,deg \leq Zenith\,angle \leq 22\,deg$ was applied beforehand ('precut'). It has to be mentioned that the cut on the zenith angle results in a sharper distribution than the total average has. The conditions for the sharpest ALPHA distribution are listed in Tab. A.3. The very best one is printed in bold.

As a conclusion of these studies we can summarize:

1. Increasing the ***exponent*** n to $n = 1.5$ results in a sharper ALPHA distribution.
2. Increasing the ***cleaning level*** to ***4.0/3.0*** also reduces mean and variance of the ALPHA distribution.
3. The ***largest*** effect is obtained by ***removing the islands*** from the image before calculating ALPHA.

Unfortunately the miss-pointing of the CT1 telescope is quite large and even after a pointing correction of the recorded data the ALPHA distribution does not have the sharpness that

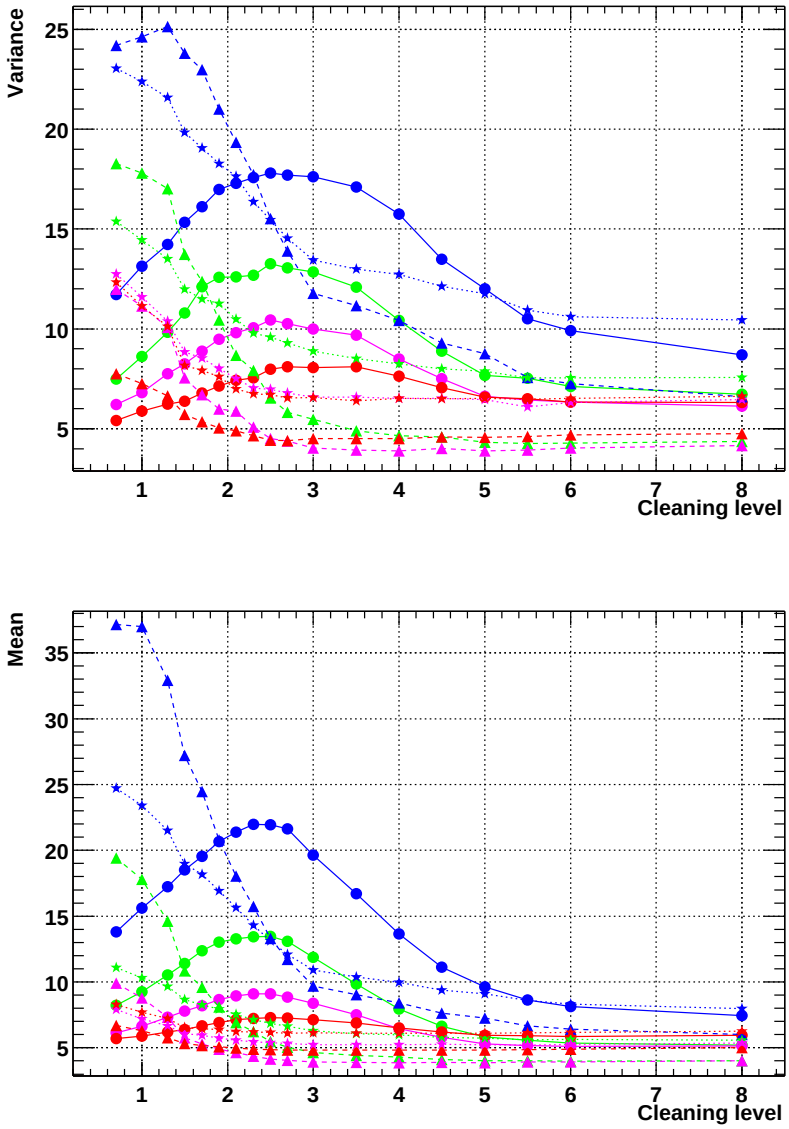


Figure A.21: This plot shows the variance and the mean of the ALPHA distribution of gamma shower images for different 'core' cleaning levels. The smaller these quantities, the better for the background reduction. The exponent n is coded in colors. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different cleaning algorithms are expressed in different marker styles and line styles. Circles and continuous line represent classic cleaning, triangles and dashed line is 'island' cleaning and stars with dotted line are 'mountain' cleaning. The sharpest ALPHA distribution is obtained for high cleaning levels above 3 and exponents of 1.5.

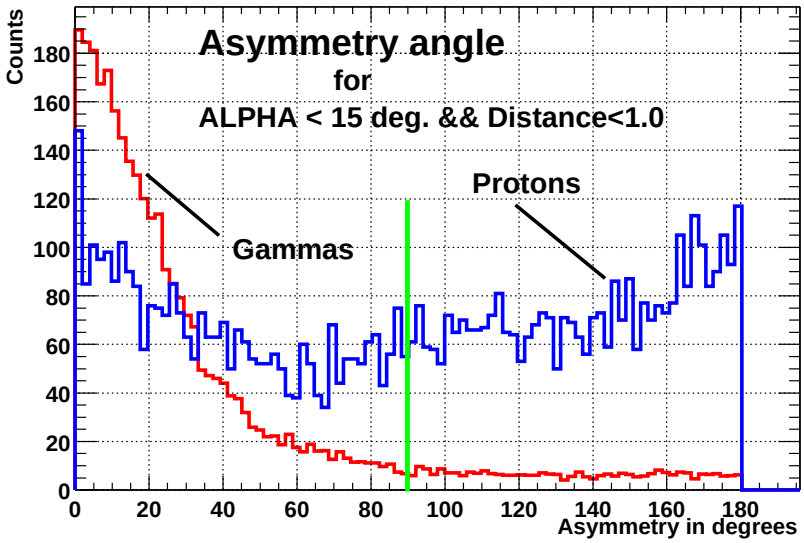


Figure A.22: The asymmetry distribution for gammas and hadrons. 0° points towards the camera center and 180° points away from the camera center. The histogram has been filled ***after a cut on ALPHA*** $< 15^\circ$ and DIST < 1.0. The cut on ALPHA has been performed to demonstrate that only in the case of γ-evens the asymmetry is correlated to ALPHA and not for hadronic events. For that reason it is possible to ***suppress in the ideal case half of the background*** (for a cut on the asymmetry angle at 90°). The hadron distribution has been scaled up for better visibility.

would be expected from MC studies. We will only be able to take little advantage of the sharper ALPHA distribution.

A.4.2 The head-tail asymmetry defined in a such way that it overcomes the finite resolution of the CT1 camera

In the chapter 1, concerning theory and detectors, it was mentioned that all shower images (including hadronic ones) have a ***head-tail asymmetry*** with respect to the direction from where they originated (Fig. A.22) due to the imaging process which involves a ***tangent*** of the viewing angle. For point sources this knowledge can be used to ***further reduce*** the background since the hadronic shower arrive randomly from all directions and the gamma showers only arrive from the center of the camera (if the center points to the source). This idea is not new and has already been used by other experiments. They defined the asymmetry in the following way [Pet97]:

$$\overrightarrow{A} = \overrightarrow{x_{HP}} - \langle \overrightarrow{x} \rangle \tag{A.69}$$

This is the difference between the coordinates of the highest pixel and the shower center of gravity (Equ. A.16). However, this definition has a problem. Since the difference between $\overrightarrow{x_{HP}}$ and $\langle \overrightarrow{x} \rangle$ and the resolution of the CT1 camera (0.25°) are of the same order of magnitude doesn't yield good results.

The definition of the asymmetry angle

Here I introduce another formalism which uses the ***weighted average*** of all pixels combined and avoids the pixel size problem. The idea behind this method is the fact that as the ***exponent*** n in Equ. A.65 ***increases***, the weighted mean (Equ. A.16) moves ***in the***

direction to the shower maximum and vice versa. Consequently, the ***derivative*** of the weighted mean with respect to the exponent leads consequently to a vector of ***steepest descent*** (as seen from the shower maximum) which characterizes an asymmetry in the image.

$$\begin{aligned} \overrightarrow{A} &= \frac{d\,\langle \overrightarrow{x} \rangle}{dn}\,|_{n=2} && \text{(A.70)} \\ &= n\left[\frac{\left(\sum_i q_i^{n-1}\overrightarrow{x}\right) - \langle \overrightarrow{x} \rangle \sum q_i^{n-1}}{\sum_i q_i^n}\right] \\ \cos\alpha &= \frac{\left\langle \overrightarrow{A}, \langle \overrightarrow{x} \rangle \right\rangle}{\left|\overrightarrow{A}\right| \left|\langle \overrightarrow{x} \rangle\right|} && \text{(A.71)} \end{aligned}$$

where α is the angle between the asymmetry vector and the vector from the shower mean to the center of the camera (similar as ALPHA). The parameter α will be called ***asymmetry angle***. The value of the angle α ranges from 0 to 180°, where 0° points towards the center and 180° points outside of the camera.

In case of γ-events (not in case of hadrons!) ALPHA is strongly ***correlated*** with the asymmetry angle α. This formula is still not perfect since many events are truncated at the camera border. With these truncations the asymmetry vector is ***no longer*** calculated correctly and it points outside of the camera instead of at the center. To avoid this problem, ***only*** the pixels within a ***radius*** of $r = Min\,(1.5 - |\overrightarrow{x_{HP}}|\,, 0.25)$ around the ***pixel with largest signal*** are included in the calculation.

Finding the best cleaning conditions and exponents n for the asymmetry

Just as with the other parameters, tests have been performed by ***systematically changing*** the values and algorithms for image cleaning and the exponents n. To determine which are the best conditions, another parameter called ***cut efficiency*** was introduced. It is the ratio of the number of events below 90° (of the asymmetry angle) to the total number of events $\varepsilon = N(Asymmetry < 90°)/N_{Tot}$.

In Fig. A.23 the ***cut efficiency*** is plotted for (recorded) ***hadronic events*** and for MC ***gamma events***. As before, circle markers represent 'classic' image cleaning, triangles 'island' image cleaning and stars 'mountain' image cleaning. The colors blue, green, pink and red represent the exponents (n=0.5, 1.0, 1.5 and 2.0).

As a conclusion it can be stated the following:

1. In case of the asymmetry a ***high cleaning level***, just as for the ALPHA distribution, is favored.

2. The best asymmetry distribution results from a ***cleaning level*** of about ***4.0/3.0*** using the ***'island' cleaning*** algorithm and an ***exponent*** of $n = 2.0$.

The resulting distribution can be seen in Fig. A.22. The asymmetry of the hadron distribution originates from events which have been partially truncated at the camera border. A cut of 0.4<DIST<1.05 , SIZE>100 and 12°<zenith angle<22° was applied beforehand.

A.4.3 The problem of truncated images due to the limited size of the CT1 camera and the leakage parameter

As mentioned above, shower images that have large impact parameters or high energies are larger than the camera's field of view so that the image is ***truncated*** at the camera border. Unfortunately, this occurs for the ***majority*** of the showers since the number of γ events

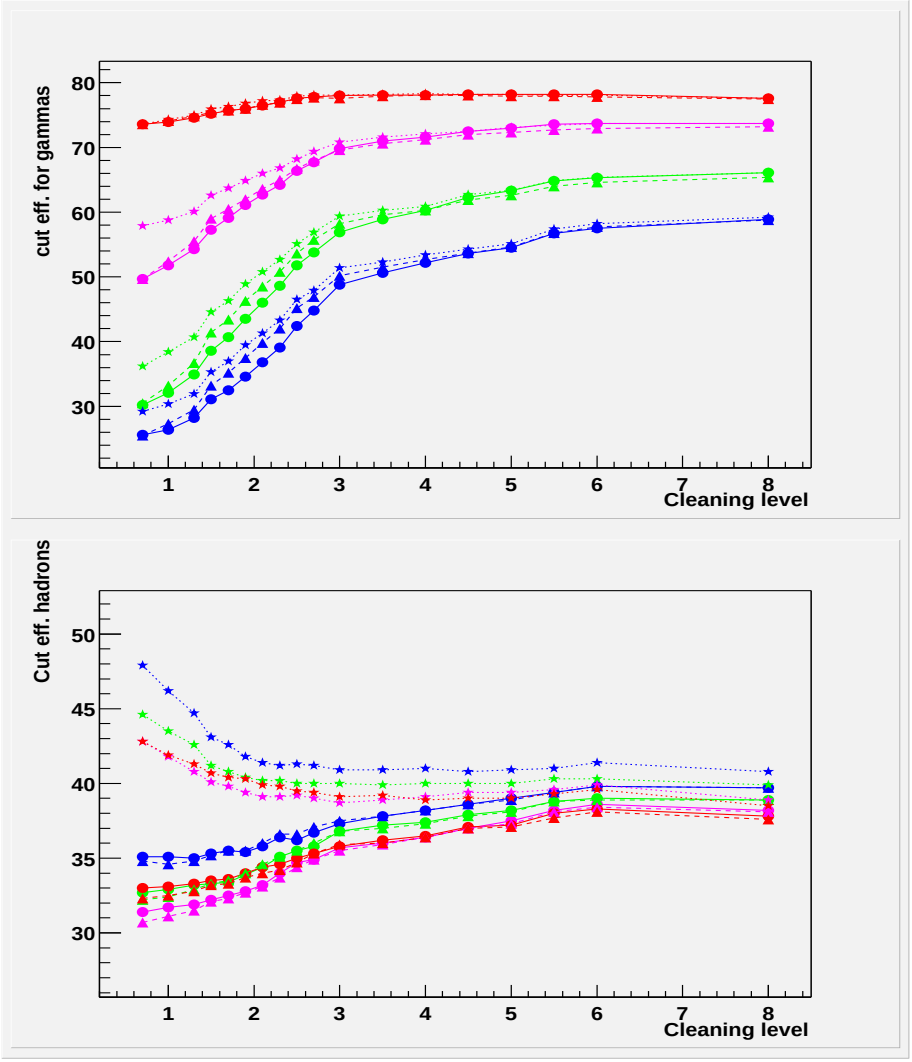


Figure A.23: In this figure the ***cut efficiency*** of gamma showers (upper plot) and hadronic showers (lower plot) for a cut of $Asymmetry < 90\,deg$ is plotted for different 'core' cleaning levels and algorithms. The ***exponent*** n is coded in colors. Blue is $n = 0.5$, green is $n = 1.0$, pink is $n = 1.5$ and red is $n = 2.0$. Different ***cleaning algorithms*** are expressed in different ***marker styles*** and ***line styles***. ***Circles and a continuous line*** represent ***classic cleaning***, ***triangles*** and a ***dashed line*** refer to ***'island'*** cleaning and ***stars with a dotted line*** represent ***'mountain'*** cleaning. The highest hadronic ***suppression*** (60 %) while retaining most gamma events (78%) is obtained for cleaning levels ***above 3*** and an ***exponent of*** n***=2.0***. The cut efficiency for gammas does never reach 100% since a cut of 0.4<DIST<1.05 , SIZE>100 and 12°<zenith angle<22° was applied beforehand.

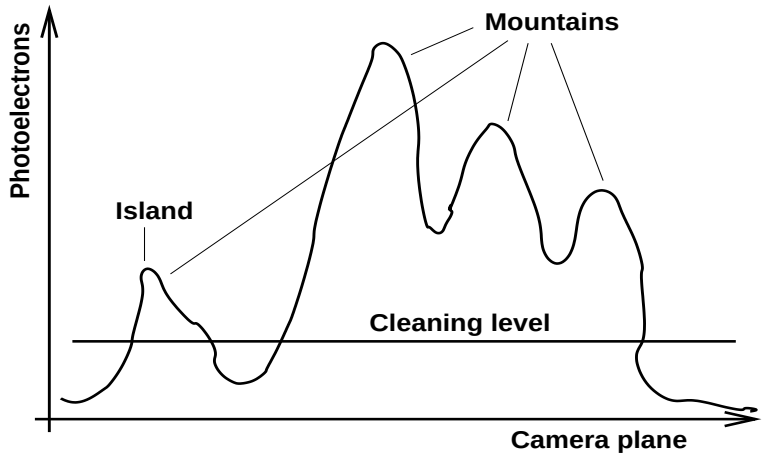


Figure A.24: The figure illustrates what is meant with the expressions ***'mountain'*** and ***'island'***. Islands can form ***after*** performing a ***'classic'*** image cleaning of a shower image, depending on the cleaning level. They are isolated clusters that have no connection to other clusters. For simplicity 'islands' are also called ***'mountains'***. In an 'island' image cleaning only islands are removed. In a 'mountain' image cleaning all the mountains and islands are ***removed*** apart of the largest mountain.

increases linearly with the impact parameter. If part of the shower is outside the camera then not all of the light will be collected and measured. This leads to an incorrect energy estimation for showers with large ***impact parameters*** and high ***energy***.

Therefore, a parameter that ***quantifies*** the truncation is needed. For this purpose, I introduce a simple parameter, called ***LEAKAGE***, defined as:

$$LEAKAGE = \frac{\sum_{Border\,pixels} q_i}{SIZE} \tag{A.72}$$

It is the ratio of the charge content of the border pixels and the total charge (i.e. SIZE). In Fig. 1.17 we see the profile of a typical shower. It is clear that by knowing the amplitude at say DIST=1.5° and also its integral from 0° to 1.5° we can estimate the content of the tail quite well. Later, I will show later that this parameter ***increases the resolution of the energy estimation***.

A.4.4 Mountains and islands in the shower image: Difference between electromagnetic showers and hadronic showers

The main difference between hadronic showers and electromagnetic showers is the larger ***WIDTH*** and larger ***fluctuations*** of hadronic showers. An hadronic shower has a core of high energy hadrons that develop ***electromagnetic subshowers***. Seen by a Cherenkov telescope a hadronic shower looks like the ***superposition*** of several, dispersed, ***low energy*** electromagnetic showers. As a consequence, the Cherenkov image of a hadronic shower has more ***'mountains'*** in the sense of Fig. A.24. The electromagnetic shower consists ideally of ***one large peak***. I tried to quantify this difference by carrying out a structure analysis, i.e. by ***counting*** the 'mountains' in the image.

The mountain finding algorithm

I performed a ***recursive cluster analysis***, by separating the image into individual clusters. Each cluster contained one mountain. The mountain search method was as follows:

- Perform a ***'classic'*** image cleaning (without removing islands).

- In a first step, find the ***peak pixels*** using a recursive algorithm. The peak pixels are the ones that have a ***larger*** content than all the neighbour pixels. Afterwards ***sort*** this list, in ascending order.
- Introduce a ***cluster object*** for each peak pixel.
- Start with the ***largest*** signal pixel (top) and search recursively ***downhill*** from the top to the ***valley*** around the mountain. Put all these pixels, ***including*** the valley pixels, into the cluster. Treat each pixel only ***once*** and ***mark*** those already used. Valley pixels are defined such: There is no unmarked neighbour pixel with lower content.
- Repeat the procedure until all the pixels have been ***assigned*** to clusters.
- Count the ***number of mountains*** (-> this is the parameter ***Number-Mountain***) and ***count the charge*** in all the mountains, except from the largest mountain. (-> this is the parameter ***Mountain-SIZE***). This parameter quantifies the charge which is ***not*** inside the largest mountain (for γ-showers ideally zero).

The two new parameters Number-Mountain and Mountain-SIZE give a good estimate of which part of the charge is found in sub-showers. This quantity is very ***different*** for gamma showers and hadronic showers.

The island finding algorithm

Before continuing with finding the best ***conditions*** for these two parameters I want to introduce another algorithm which is used for the ***'island' image cleaning***. It is similar to the cluster/mountain analysis and I called it the 'island' analysis. An 'island' is a cluster that is isolated in a cleaned camera. A shower image (after 'classic' cleaning) usually consists of one main island (the main shower image) and some additional islands which are much ***smaller.***

The algorithm does the following:

- Perform a ***'classic'*** image cleaning
- Loop over the pixel list and call a ***recursive*** algorithm finding the neighbour pixels for each pixel. Continue the recursion with the neighbour pixels. ***Mark*** each seen neighbour pixel and stop at the ***border*** of the island. Treat only unmarked pixels. Introduce a ***cluster object*** for each new unmarked pixel and assign all the pixels of one ***island*** to this cluster.
- Count the ***number of islands*** and calculate the ***charge*** of each island.

The cluster list obtained by the island or mountain analysis can be used for ***image cleaning***, leaving only the largest island or mountain. I call these cleaning algorithms ***'island'*** and ***'mountain'*** cleaning.

Gamma/hadron separation

In this work only the mountain analysis is used for gamma/hadron separation purposes. As before, a systematic procedure was used to check the best ***conditions*** in terms of cleaning level, to provide fore for best separation of gammas and hadrons. The parameters 'separation' and 'overlap' are shown in Fig. A.25 and Fig. A.26.

The results are displayed in Tab. A.4. In these two cases the smallest overlap does not coincide with best separation because of completely non-Gaussian distributions. The value that was chosen as ***optimal cleaning level was 3.5/2.7***. (The cleaning level consists of two numbers. A 'core' limit and a 'border' limit.).

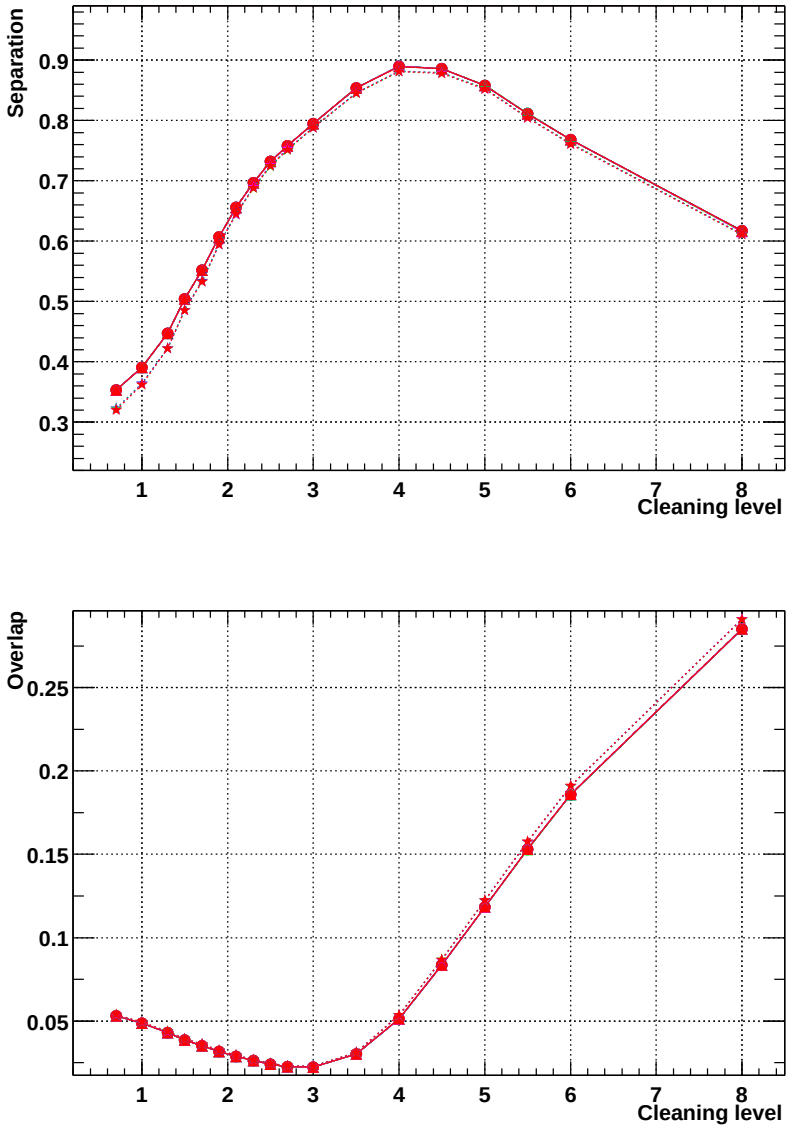


Figure A.25: The separation (above) and the overlap (below) for the parameter ***Mountain-SIZE***. Optimal values for these two parameters are found for ***'core'*** cleaning levels between ***3 and 4***. The optimal values for the cleaning level for both parameters don't ***coincide***, therefore the average of a level of ***3.5*** has been chosen.

Parameter	Cleaning level	Separation	Overlap
MultMountain	3.5/2.7	0.983	0.1083
MountainPhe	3.5/2.7	0.890	0.0512

Table A.4: The table presents the cleaning levels which provide the best separation and the smallest overlap. The cleaning level consists of two numbers. A 'core' limit and a 'border' limit.

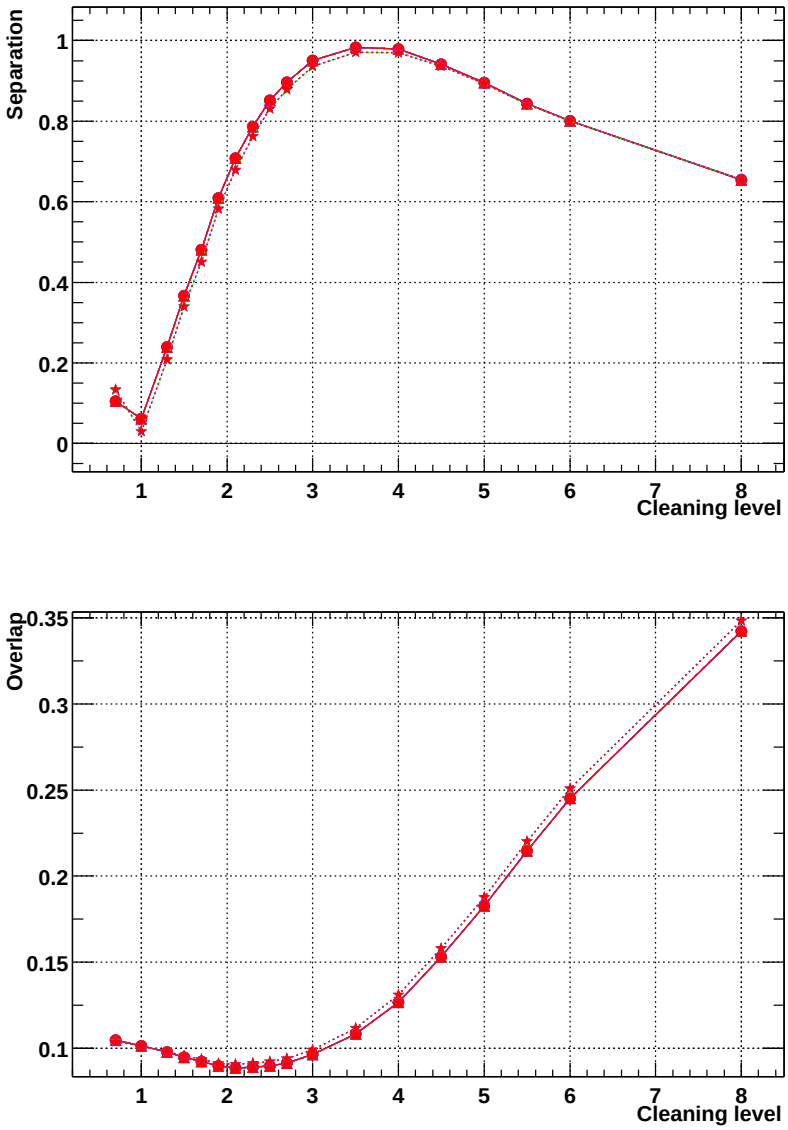


Figure A.26: The separation (above) and the overlap (below) for the parameter ***Number-Mountain***. The parameters separation/overlap have optimal values for ***'core'*** cleaning levels between ***2 and 4***. They cleaning levels in the lower and upper plot . An intermediate value of a level of ***3.5*** was chosen.

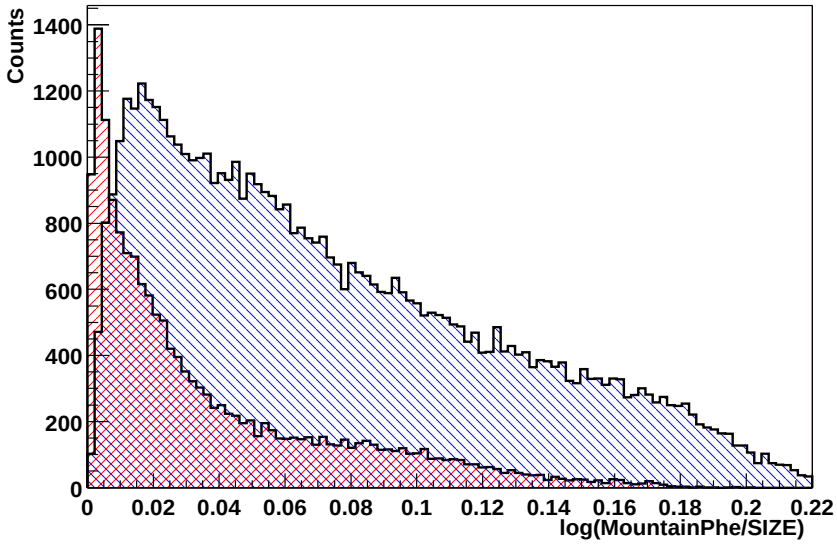


Figure A.27: Distribution of the Mountain-SIZE parameter, plotted as the logarithm of its ratio of Mountain-SIZE with SIZE for better illustration, for gammas (red) and for hadrons (blue). This distribution is not useful for hard cuts, but it improves the separation if used as a input in the LDA.

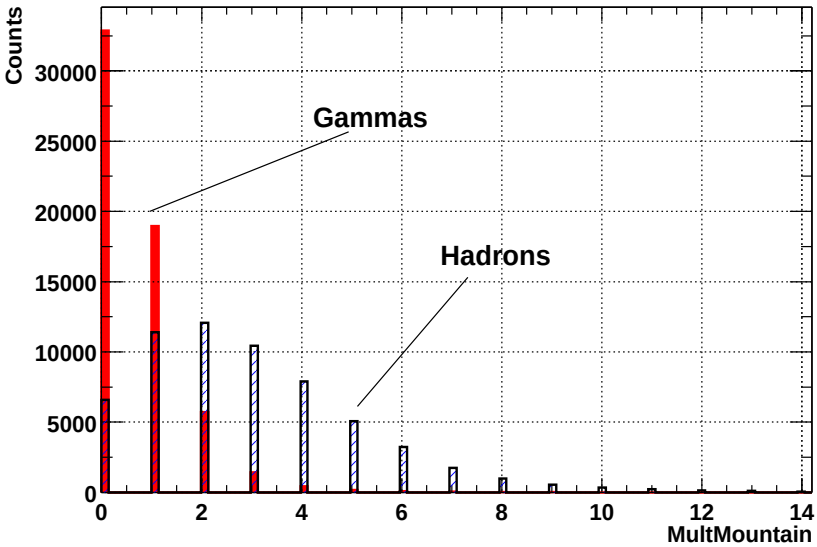


Figure A.28: Distribution of the Number-Mountain parameter for gammas (red) and for hadrons (blue). A hard cut on Number-Mountain < 2 removes a sizable portion of the background while leaving most γ's.

Parameter	Cleaning procedure	Level	Exponent
ALPHA	'Island'	4.0/3.0	1.5
WIDTH	'Island'	1.3/0.3	1.5
LENGTH	'Island'	1.3/0.3	1.0
CONC	'Island'	1.3/0.3	/
DIST	'island'	1.3/0.3	1.5
Asymmetry	'classic'	4.0/3.0	2.0
NumberMountain	'classic'	3.5/2.7	/
MountainSize	'classic'	3.5/2.7	/

Table A.5: The table summarizes the optimal cleaning levels for the different image parameters.

The ***Number-Mountain*** distribution and ***Mountain-SIZE*** distribution are plotted for both, gammas and hadrons in Fig. A.28 and Fig. A.27.

The Number-Mountain distributions for gammas and hadrons are quite ***well separated***, but the one of the Mountain-SIZE parameter shows significant overlap and it is ***not useful for hard cuts***. However, it still improves the separation when used as an ***additional input*** to the LDA together with other parameters. This will be shown in the next section.

A.4.5 Improvement in gamma/hadron separation

As all of the new parameters have been now presented, we can proceed to check whether they give an improvement in discrimination of the total image dataset. The tool for this procedure is the ***LDA***. More parameters will be introduced, step by step, to see if an ***improvement*** can be achieved.

Improvement of the gamma/hadron separation by including the new parameters in the LDA

Now we include the new parameters, that were defined above, in the analysis and see if this increases the discrimination power. The new parameters are

- Parameters with ***weights*** (***exponents*** n) on q_i.
- Parameters obtained by ***different image cleaning*** procedures.
- ***Asymmetry***
- ***Number-Mountain***
- ***Mountain-SIZE***

Tab. summarizes the optimal cleaning levels for the different image parameters as obtained in the studies above. The image parameters of ***three datasets*** were calculated accordingly. These are the ***MC gamma dataset***, a ***(recorded) OFF*** data set and the ***Mkn 421 dataset*** (flares in 2001, 167 hours observation time).

The ***stepwise improvement*** is shown in ***four*** steps:

1. First the parameters ***WIDTH***, ***LENGTH***, ***CONC***, ***zenith angle*** $\cos\theta$, ***SIZE*** and some ***higher orders*** are included in the LDA input parameter list, as described in section A.3.5. The gives us the first nine inputs from input[0]...Input[8]:

$$\begin{aligned} Input[0] &= WIDTH \\ Input[1] &= LENGTH \\ Input[2] &= CONC \end{aligned} \tag{A.73}$$

$$
\begin{aligned}
Input[3] &= SIZE \\
Input[4] &= \cos\theta \\
Input[5] &= WIDTH^2 \qquad \text{(A.74)} \\
Input[6] &= LENGTH^2 \qquad \text{(A.75)} \\
Input[7] &= CONC^2 \qquad \text{(A.76)} \\
Input[8] &= \frac{WIDTH * LENGTH}{SIZE} \qquad \text{(A.77)}
\end{aligned}
$$

2. Then the ***asymmetry angle*** α of Equ. A.71 and the its cosine were included.

$$
\begin{aligned}
Input[9] &= \alpha \qquad \text{(A.78)} \\
Input[10] &= \cos\alpha
\end{aligned}
$$

3. After this the mountain classification parameters ***Number-Mountain***, ***Mountain-SIZE*** and ***Mountain-Size/SIZE*** were included.

$$
\begin{aligned}
Input[11] &= NumberMountain \qquad \text{(A.79)} \\
Input[12] &= MountainSize \\
Input[13] &= MountainSize/SIZE
\end{aligned}
$$

4. In the end the ***remaining rest*** was included, which give still ***small corrections*** and improve separation

$$
\begin{aligned}
Input[14] &= Leakage \qquad \text{(A.80)} \\
Input[15] &= Leakage^2 \\
Input[16] &= Distance \\
Input[17] &= AveragePedestalSigma
\end{aligned}
$$

At each step the LDA was trained with the MC-gamma-dataset and the OFF-dataset which gives us as a result the ***discriminating power***. Then the trained LDA was applied to the dataset of ***Mkn 421*** (flares in 2001, 167 hours observation time). This yielded the number of ***excess events***, the ***background*** and the resulting ***significance (Li/Ma)***. The following static fix quality selection cut has been applied before the LDA procedure: $0.5 \leq DIST \leq 1.05$, $SIZE > 60\,PhE$, $Zenith\,angle < 50\,deg$, $WIDTH > 0.25$ and $LENGTH > 0.4$.

The results are presented in the Tab. A.6. In the first row the result of the ***simple static cut*** is shown. The next two rows show the results using classic image parameters with 1. ***linear input*** and 2. ***higher order input.*** The last ***four*** lines show the stepwise improvement by using parameters with weights and new parameters as listed before.

The final ALPHA plot for the ***Mkn 421*** test sample (167 hours of observation) can be seen in Fig. A.29. An OFF-data sample with lower statistics is also plotted in the same figure to ***demonstrate*** that the background estimation by fitting a polynomial is in agreement with the OFF-data distribution. The ALPHA distribution is wider than the previous using classical image parameters because much more ***additional low energy events*** (which have an unsharper ALPHA distribution) were found. The final result are about 16000 +- 250 excess events with a significance (Li/Ma) of about 83 sigma.

Fig. A.30 shows the output of the LDA. Fig. A.31 displays the cut efficiency of the ***final*** version plotted against zenith angle and energy. The low energy cut-efficiency improved quite significantly and it's almost flat in zenith angle response. The average cut efficiency (after trigger and 'precut') is now as high as 80 %. To ***avoid confusion*** it has to be mentioned that the ***precut on DIST*** is ***dependent*** on the exponent n=1.5. For this

Step	Disk. power	Excess events	Backg.	Sig.
static cuts	/	5942.9+-134	3869	47.9
classic parameters linear	0.678	11389+-189	9681	60.8
classic parameters quad.	0.736	12592+-204	9158	67.3
1. Parameters quad.	0.755	14486+-256	11760	70.4
2. Plus asymmetry	0.766	16079+265	10886	77.8
3. Plus mountains	0.769	15636+-242	9275	80.0
4. Plus rest	0.777	16182+-245	9060	82.8

Table A.6: The table presents the stepwise increase of discriminating power, excess events and significance while reducing the background by including the new image parameters. The first line contains the results of the static cuts. The next two lines show the numbers for classic HILLAS parameters using linear and higher order inputs in the LDA. The next four steps show the increase in discrimination of hadrons and gammas by using image parameters with weights and including new parameters. In the first step the parameters ***WIDTH***, ***LENGTH***, ***CONC***, ***zenith angle*** $\cos\theta$, ***SIZE*** and some ***higher orders*** are used as inputs for the LDA. In the next three step the list is expanded by the asymmetry, mountain classifications and in the end by parameters that correct for correlations.

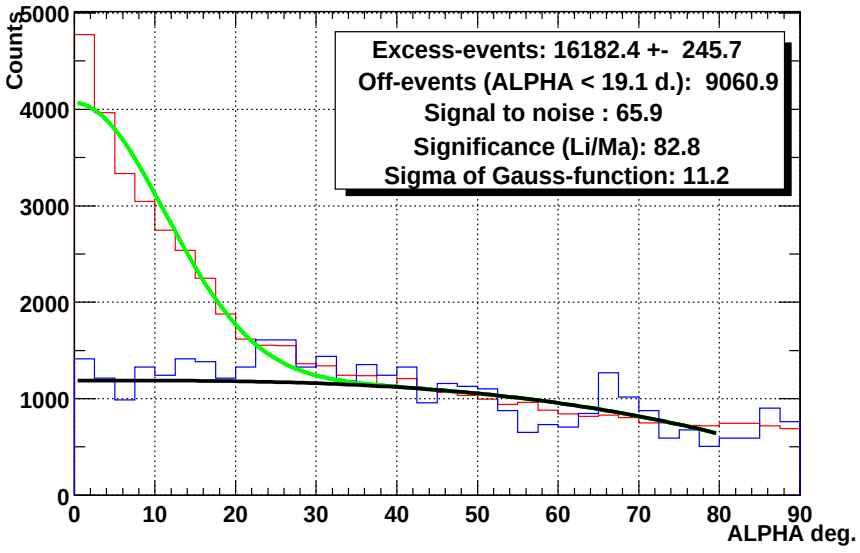


Figure A.29: The ALPHA plot of the final LDA cut. The plot shows the final result with the maximum significance of 82.8 sigma that achieved for the dataset of Mkn 421 of 167 hours of observation time.

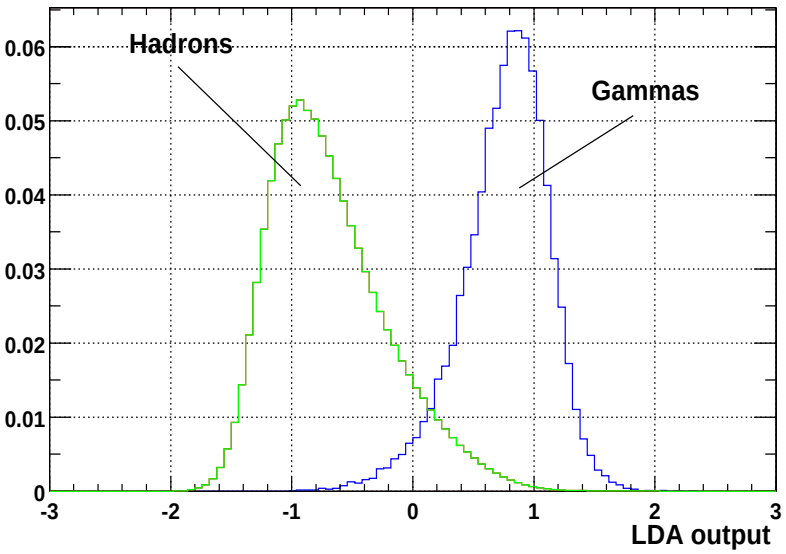


Figure A.30: The histogram shows the output distributions (normalized to one) of the LDA for (recorded) hadrons and (MC) gammas after introducing all possible improvements.

reason the average cut efficiency after precut here ***cannot*** be compared directly with the value obtained at the beginning when using classical HILLAS parameters (the average cut efficiency after precut was around 70 %, but the sample was smaller due to a different definition of DIST with $n = 1.0$). Apparently the cut efficiency increased only by 13 % while the excess events increased by 30 %, which is ***no contradiction***. The precut on DIST has always been chosen to result in a ***maximally large*** significance.

Fig. A.32 shows the effective area according to the LDA selection cut (which was optimized for maximal significance). In the calculation of the effective areas the telescope trigger efficiency is also included. This is why the collection area for energies below 1 TeV is rather small while the cut efficiency is still very high, approximately 60%. The effective areas for the ***lowest zenith angle*** are ***above 55 000*** m^2 (below 1 TeV). They reach almost ***95 000*** m^2 for the largest zenith angle of ***45°***.

A.4.6 Conclusion about the introduction of new image parameters and new image cleaning algorithms

It has been shown that the introduction of ***exponents*** on the charge in each pixel gave clear improvements of the ***separation power*** of the image parameter set in case of the parameters ***ALPHA***, ***WIDTH*** and ***LENGTH*** while the introduction of weights ***depending on the noise*** in each pixel ***did not*** improve the discrimination.

New image parameters like the ***asymmetry angle*** of the a shower, ***mountains*** and leakage where clearly able to further reduce the background.

A ***significant improvement*** came from the usage of a ***different cleaning algorithm***, the ***'island'*** cleaning which allowed to ***decrease*** the cleaning level and thus keep more information for the separation. This improves the separation for small energy events.

Altogether (also by using the LDA as separation tool) it was possible to ***double the significance*** of the detected signal and to almost triple the amount of ***signal (excess) events*** (always at maximum significance).

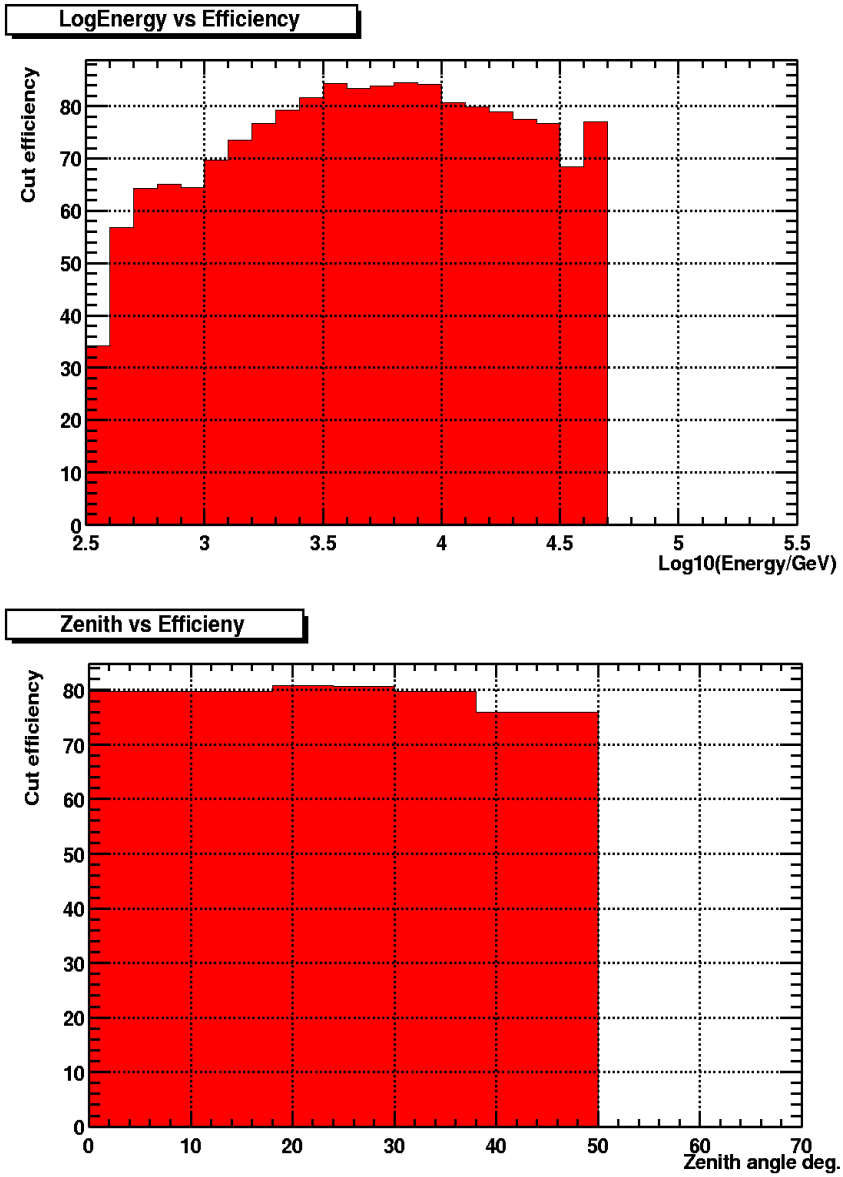


Figure A.31: The upper plot shows the cut-efficiency vs log(Energy/GeV) and the lower plot shows the γ cut efficiency vs zenith angle (in degrees). The LDA has a quite good cut efficiency for energies smaller than 1 TeV. Above 500 GeV the cut efficiency is still better than 60 % to 65 %. Above 10 TeV the efficiency decreases slightly because very big showers (high energy) are always truncated at the camera border. The cut efficiency is measured after the trigger and after applying a filter cut of $0.5 \leq DIST \leq 1.05$, $SIZE > 60\,PhE$, $Zenith\,angle < 50\,deg$ and a two-next-neighbour software trigger. The lower plot shows an average cut efficiency of 80 % for a cut that results in maximal significance. The DIST pre-cut is dependent on the exponent n. For this reason the improvement shown here seems smaller than it actually is, compared to Fig. A.12.

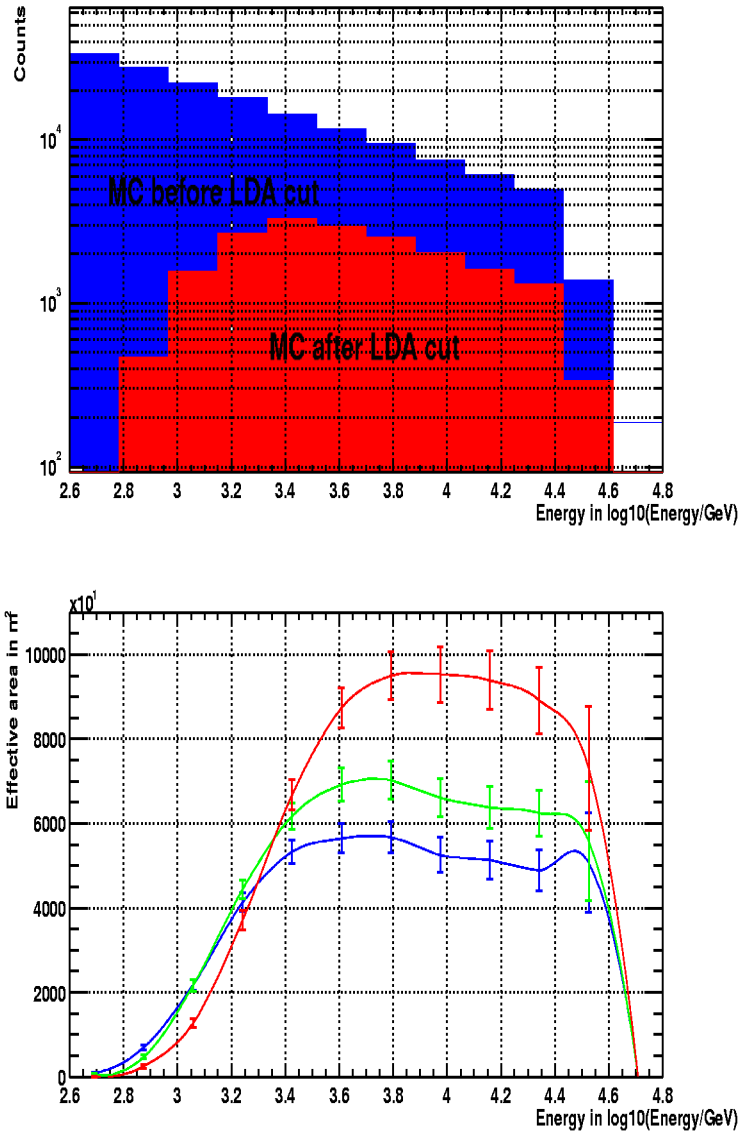


Figure A.32: The lower plot shows the effective areas calculated for the LDA selection cut using a power law spectrum with a spectral index of $\alpha = 2.9$ for three zenith angles. The blue curve is for a 12°, the green curve for a 32° and the red curve for a 45° zenith angle. The upper plot shows the simulated MC events before cut (the blue curve) and the triggered events after LDA cut (the red curve).

A.5 Energy estimation by the least squares method

Now, that we established our separation methods, we would like to use the selected events to calculate the physical quantities that we are interested in. The first step is to estimate the energy from the image parameters and to determine the energy resolution that we can get for CT1.

Here I use a very simple method based on image parameters and the ***method of linear least squares*** to ***estimate the energy***. As we will see, the formalism is quite similar to that of the LDA.

The method of linear least squares and its application

We have a set of parameters for each event i

$$\overrightarrow{p^i} = \begin{pmatrix} 1 \\ p_1^i \\ \vdots \\ p_n^i \end{pmatrix} \equiv \mathbf{P}_k^i \tag{A.81}$$

which parameterize the image. The 'ONE' in the first row is important as it is needed for the offset. We wish to find an estimator E_{est}^i for the energy E^i of each event i in such a way that the ***sum of differences***

$$S = \sum (E_{est}^i - E^i)^2 \tag{A.82}$$

with the linear ansatz

$$E_{est}^i = \mathbf{P}_k^i w^k \tag{A.83}$$

being ***minimal.*** This is the case if

$$\begin{aligned} \frac{\partial S}{\partial p_j} &= 2\left(\mathbf{P}^T\right)_i^j \left(\mathbf{P}_k^i w^k - E^i\right) = 0 && \text{(A.84)} \\ \left[\left(\mathbf{P}^T\right)_i^j \mathbf{P}_k^i\right] w^k &= \left(\mathbf{P}^T\right)_i^j E^i && \text{(A.85)} \\ w^k &= \left[\left(\mathbf{P}^T\right)_i^j \mathbf{P}_k^i\right]^{-1} \left(\mathbf{P}^T\right)_i^j E^i && \text{(A.86)} \end{aligned}$$

This procedure is called ***linear least squares*** and can be ***solved analytically***. (It is almost identical to the calculation of the weights of the LDA. In case of the LDA the E^i correspond to the different groups. For the case of 2 groups it becomes $E^i = -1$ for gammas and $E^i = +1$ for protons. The 'ONE' above corresponds to the subtraction of the average output given by Equ. A.61). It only remains to calculate the ***weights*** w^k using Equ. A.86 which involves the calculation of an ***inverse matrix***. The inverse matrix calculation was done by the Gaussian elimination algorithm using a precision of 24 digits.

The following parameters have been used to estimate the (log) energy.

$$\begin{aligned} p_0 &= \log(SIZE) && \text{(A.87)} \\ p_1 &= \sqrt{\log(SIZE)} \\ p_2 &= \log^2(SIZE) \\ p_3 &= \frac{1}{\cos(ZenithAngle)} \\ p_4 &= \log(LENGTH) \\ p_5 &= DIST \end{aligned}$$

The ***main estimator is SIZE*** (and expressions of it) which is corrected with the ***cosine of the zenith angle*** (shower distance changes with cosine). The other parameters are introduced to obtain ***small corrections***. SIZE is dependent on the impact parameter (-> estimator DIST). For the training a MC-gamma data sample was used. A preselection cut of $ZenithAngle \leq 50\,deg$, $SIZE \geq 60\,PhE$ and $0.5\,deg \leq DIST \leq 1.05\,deg$ was applied beforehand.

Additional energy dependent ***weights*** $E^{1.5}$ on each event have been introduced in the method to ***correct*** for the fact that the MC was produced with a steep spectrum of $E^{-1.5}$ while we want that high energy events are treated with the ***same priority*** as low energy events. The whole procedure is calculated in the log-scale. Equ. A.85 and Equ. A.86 become:

$$\left[\left(\mathbf{P}^T\right)_i^j \mathbf{P}_k^i \left(E^i\right)^{1.5}\right] w^k = \left(\mathbf{P}^T\right)_i^j \left(E^i\right)^{1.5} \log\left(E^i\right) \quad \text{(A.88)}$$

$$w^k = \left[\left(\mathbf{P}^T\right)_i^j \mathbf{P}_k^i \left(E^i\right)^{1.5}\right]^{-1} \times \left(\mathbf{P}^T\right)_i^j \left(E^i\right)^{1.5} \log\left(E^i\right) \quad \text{(A.89)}$$

The result can be seen in Fig. A.33. SIZE was introduced in parallel with several exponents (1, 0.5 and 2). This improves the ***linearity*** of the estimation.

Improvement of the energy resolution by including the LEAKAGE parameter into the Least square fit

Now we will see how the resolution and the shape can still be improved by introducing a ***new parameter called LEAKAGE*** into the least square fit. It is introduced as an ***additional input*** to the least squares method of Equ. A.87:

$$P_6 = LEAKAGE \quad \text{(A.90)}$$

$$p_7 = LEAKAGE^2 \quad \text{(A.91)}$$

The result can be seen in Fig. A.34. The distribution becomes more ***narrow and Gaussian***. The distribution is not perfectly linear and exhibits a slight curve. Without unfolding the spectrum this would introduce a systematic error. The unfolding procedure corrects for this and in fact any shape of the energy estimation would be translated correctly as long as it is ***monotonically.***

A.5.1 Conclusion about the energy estimation using the linear least squares method

Using an algorithm like ***linear least squares*** to estimate the energy has the main advantage, as the ***LDA***, that the problem is ***analytically solvable*** (reproducibility, no dependence on initial values) unlike other methods like neural nets or nonlinear equations which need an ***iterative*** optimization. A matrix of the size ***8x8*** has to be inverted (best done with 24 digits precision or more). The applied energy estimation is done by calculation of a simple polynomial which is fast. The energy estimation is linear (above 1 TeV) and Gaussian distributed. The introduction of the LEAKAGE parameter ***increases*** the resolution a little and results in an output which is more Gaussian.

A.6 Mispointing of the telescope and its correction

Unfortunately the hour angle axis of the CT1 telescope is ***not perfectly aligned*** with the earths axis of rotation and in addition, the telescope structure bends slightly under its

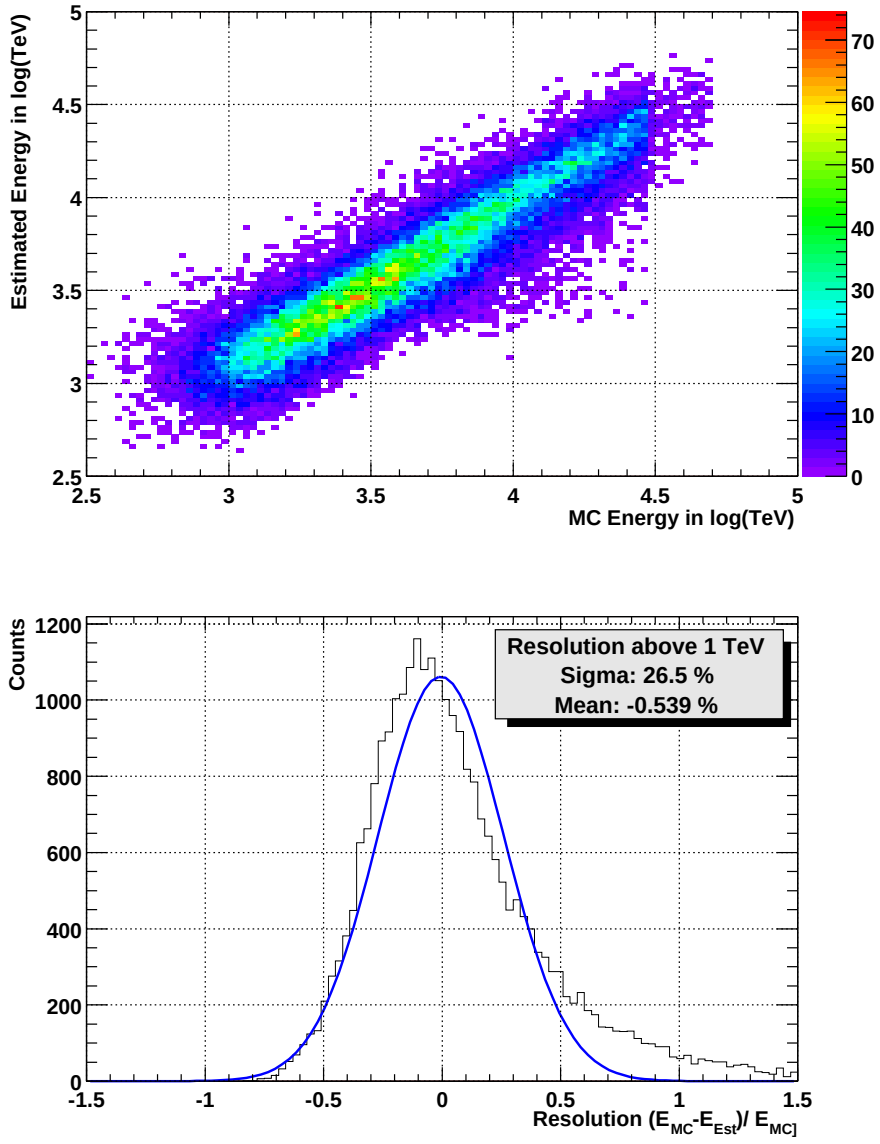


Figure A.33: Energy estimation and its ***resolution*** with the LEAKAGE parameter. For energies above 10 TeV (the upper plot) it can been seen that the effect of truncated images at the camera border disturbs the energy estimation. The lower plots shows a rather asymmetric shape which is not Gaussian.

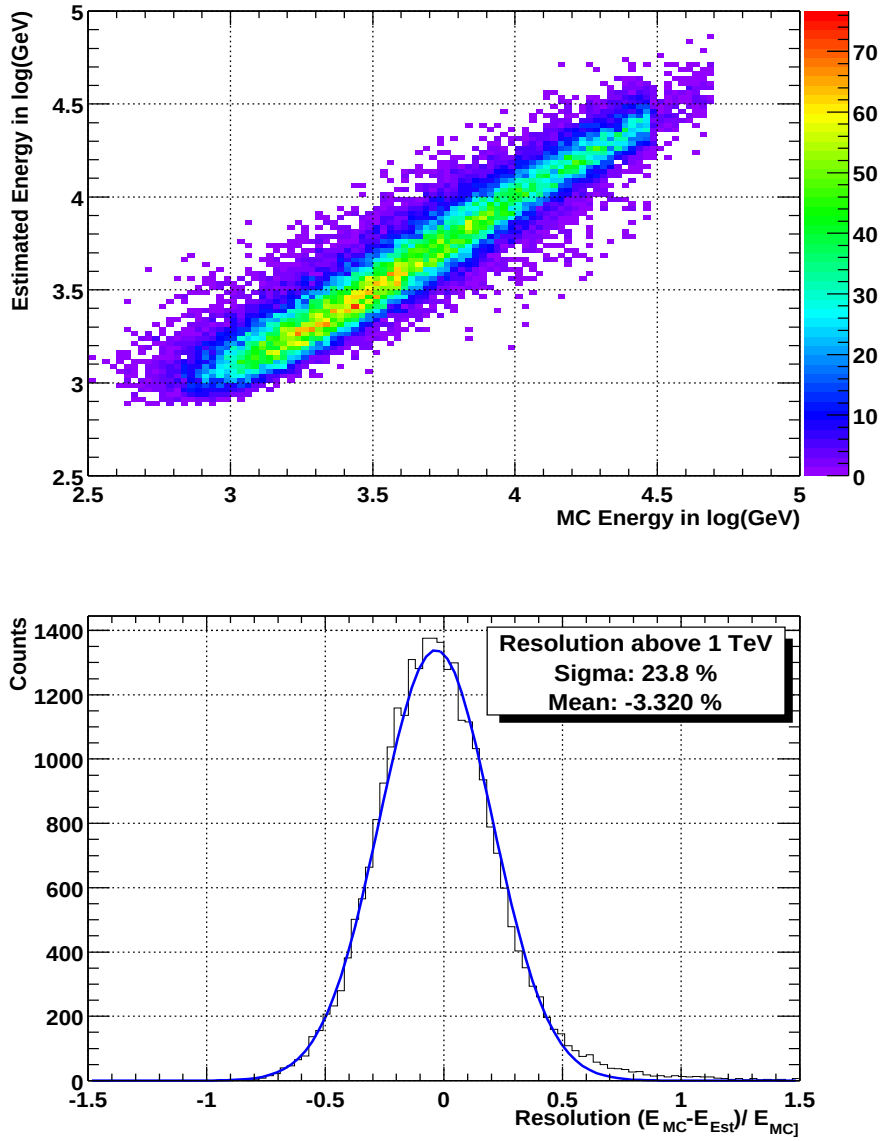


Figure A.34: The shape of the estimated energy distribution ***after*** introduction of the ***LEAKAGE*** parameter. The energy estimation becomes sharper and more precise as can be observed in the upper plot. Above 1 TeV the estimation is very linear. The lower plot demonstrates a symmetric Gaussian distribution with a variance of approximately 24 %.

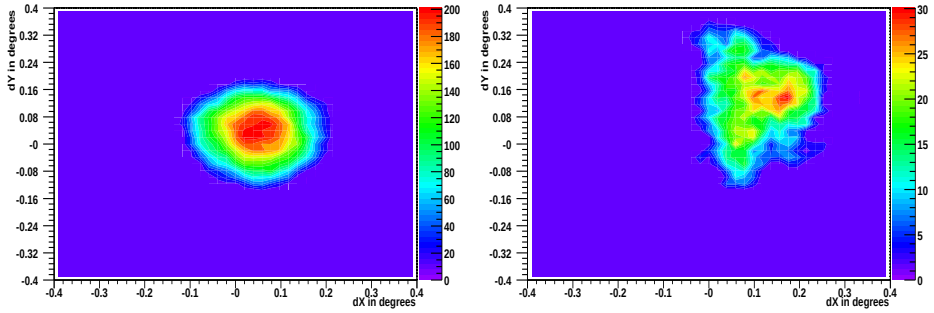


Figure A.35: A false source plot for a) a small zenith angle (15°), used to determine the mispointing of the source Mkn 421. The center of the source appears as a round spot. b) False source plot for a high zenith angle (45°). The source seems not to be focused very well in one point. This effect could be partly due to statistics but a mispointing correction on ***run basis*** results in a ***sharper*** ALPHA distribution, suggesting that for different runs with the ***same*** hour angle (= 'signed' zenith angle) the mispointing is different.

weight. This introduces in a ***mispointing*** of up to ***0.15***° and has a strong effect on the ***sharpness*** of the ALPHA distribution. Monte Carlo studies show that the standard deviation of the ALPHA peak should be approximately 6° for small zenith angles and up to 7.5° for high zenith angles of 45° (showers with higher zenith angles show a wider alpha distribution due to a smaller image in the camera resulting in a worse determination of the shower axis). With Mispointing, these values ***increase*** to a standard deviation of up to ***15***° for high zenith angles. If in such a scenario a fixed alpha cut of only 12° is applied, it is easy to imagine that the measured flux will suddenly depend strongly on the zenith angle. This introduces a ***large systematic error*** in lightcurve calculation. Thus, a pointing correction is mandatory.

Up to now, so-called point-runs have been performed where the telescope systematically scans a star (whose coordinates are known) in small steps. The DC current information of the pixels allows a precise determination of the mispointing at a given hour angle and declination. This procedure can be repeated for many hour angles and declinations. A correction based on values obtained from this procedure improved the Mispointing but did not yield completely satisfactory results and a ***different strategy*** has been tried here in this thesis.

The false source plot method

A mispointing correction can also be obtained by ***only*** using the measured data of ***Mkn 421.*** The construction of a false source plot is a simple method to find the ***real position*** of a source in the camera. Since the coordinates for Mkn 421 are known the ***mispointing*** can be calculated.

The algorithm proceeds as follows: The camera center is ***moved*** artificially (and ALPHA is recalculated accordingly) in a grid around the camera center. A ***cut*** of $ALPHA_{new}$ <8° is applied to the new ALPHA value for each position and the remaining events are ***filled*** in a 2D histogram which binned ***according*** to the grid. After the histogram has been filled, a tail-cut at the half maximum is performed by subtracting the half maximum from each bin. Bins below zero are set to zero. The position of the source is obtained by calculating the ***mean*** of the 2D histogram.

The new ALPHA is calculated from the old (signed) ALPHA as follows (see Fig. A.36):

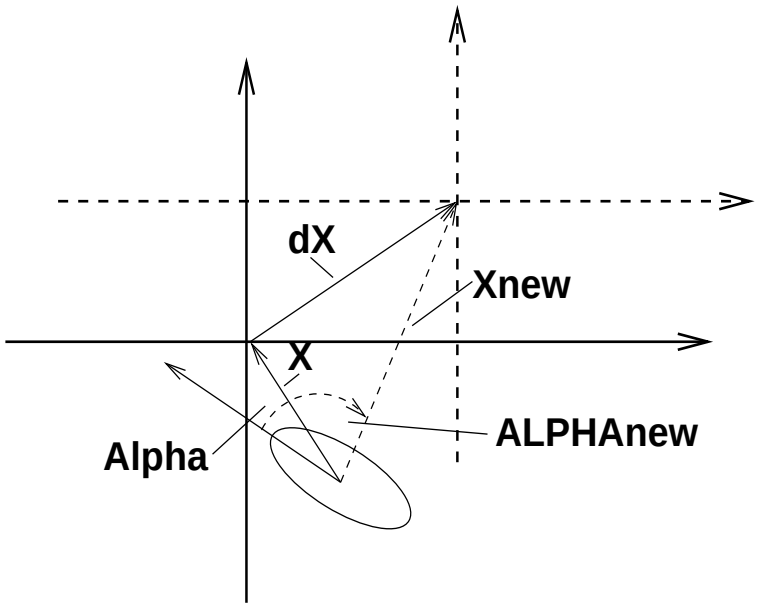


Figure A.36: Schematic of the shift of the old coordinate system and the calculation of the new ALPHA.

$$\overrightarrow{Alpha_{old}} = \begin{pmatrix} \cos\alpha_{old} & -\sin\alpha_{old} \\ \sin\alpha_{old} & \cos\alpha_{old} \end{pmatrix} \overrightarrow{\langle X_{old}\rangle} \quad \text{(A.92)}$$

$$\overrightarrow{\langle X_{new}\rangle} = \overrightarrow{\langle X_{old}\rangle} + \overrightarrow{dX} \quad \text{(A.93)}$$

$$ALPHA_{new} = \arccos \frac{\overrightarrow{Alpha_{old}} \cdot \overrightarrow{\langle X_{new}\rangle}}{\left|\overrightarrow{Alpha_{old}}\right| \cdot \left|\overrightarrow{\langle X_{new}\rangle}\right|} \quad \text{(A.94)}$$

First a vector pointing in the direction of the shower $\overrightarrow{Alpha_{old}}$ is obtained by rotating the mean position vector $\overrightarrow{\langle X_{old}\rangle}$ of the image by ALPHA ($ALPHA_{old}$). The new position vector $\overrightarrow{\langle X_{new}\rangle}$ is calculated by translating the old center with $\overrightarrow{dX}$. The new ALPHA $ALPHA_{new}$ (unsigned) is obtained from the scalar product of the old direction vector and the new position vector.

Binning the data in (signed) zenith angles

The mispointing is assumed to be a function of the hour angle and the inclination. The inclination remains ***constant*** for astronomical objects. Since the parameter hour angle was not available in the dataset a binning in ***signed*** zenith angle was performed instead (positive sign for azimuth angles larger than 180°, and negative sign for azimuth angles smaller then 180°). This binning is ***equivalent*** to binning in shaft-encoder values (and to the hour angle) of the telescope.

For ***each*** signed zenith angle bin a ***false source plot*** is calculated. By examining the false source plots one can see that for high zenith angles especially the pointing is ***not sharp*** in one spot (see Fig. A.35). This effect could be partly due to ***statistics***, but as we will see a mispointing correction on run basis results in a sharper ALPHA distribution, suggesting that for different runs with identical hour angle (equivalent to 'signed' zenith angle) the mispointing is different.

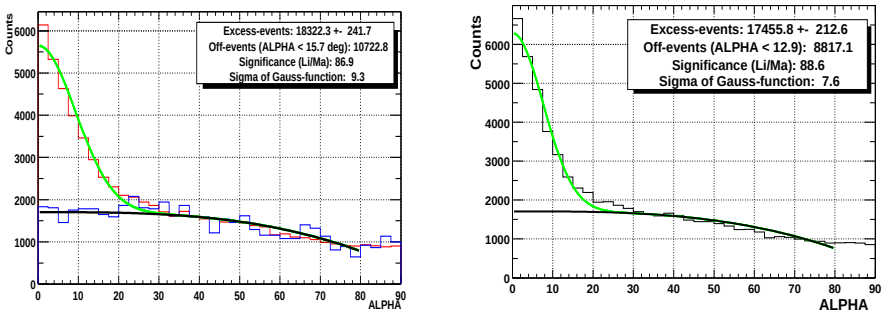


Figure A.37: a) The ALPHA distribution before applying the pointing correction. The sigma of the ALPHA peak is 9.3°. The Off-data distribution (much lower statistics) has been included into the plot for illustration. b) The ALPHA distribution after applying the pointing correction procedure. The distribution gets significantly sharper. The average sigma of the (full) Gaussian fit curve is only 7.6° . The tail of the ALPHA distribution between 20° and 30° does not fit very well to the Gaussian function. This is caused by a still imperfect mispointing correction. The high zenith angles have a wider ALPHA distribution. To avoid confusion it has to be stated that he data used here is the full dataset of Mkn 421 of 250 hours of observation time. This dataset contains more signal than the test data set used for the development of an improved analysis.

Binning in runs

To circumvent this problem, in addition a pointing correction for each run was introduced. A false source plot was then calculated for ***each run***. The correction in run bins is ***only possible*** if there is ***enough*** signal in that specific run data (the problem of a flaring source). The procedure applied was the following: If the signal in the run data was ***too small*** (smaller than ***5 sigma*** of the background fluctuations), the pointing correction on ***zenith angle bin*** basis was applied (the zenith angle bins have a ***much larger statistics and*** there is ***no danger*** of optimizing on fluctuations). If the signal was large enough during that run (larger than ***5 sigma***) then the mispointing ***based on single run*** bins is chosen. The mispointing correction on a ***run to run basis*** improves the sharpness of the ALPHA distribution significantly when compared to a mispointing correction on the basis of ***signed zenith angles*** only. This seems to point out that the mispointing for the same hour angle but on different days is different.

Results of the mispointing correction

Tab. A.7 shows the results ***after*** the pointing correction. Listed are the ***variances*** of the alpha distribution for each signed zenith angle bin. The total standard deviation was improved from ***9.3***° (before correction) to ***7.6***° (after correction). The correction on the basis of signed zenith angle bins only gave a standard deviation of approximately ***8.3***°. The table clearly shows how the alpha distribution becomes wider with ***increasing*** zenith angle. For ***all*** zenith angle bins the alpha distribution is ***wider*** than the corresponding distribution in the ***MC*** which points to still ***imperfect*** mispointing correction.

Conclusion

In conclusion, it can be stated that the result is still ***not perfect***. The variances of the alpha peak should be ***more narrow***. However, they are ***good enough*** to ensure a reasonable (without systematics coming from the ALPHA distribution) lightcurve. In the following chapters all of the cuts on the alpha distribution are performed very high at 18°. In this

Signed zenith angle	Sigma of alpha distribution
-(45-50)	10.3 ± 0.15
-(35-45)	8.9 ± 0.10
-(25-35)	6.7 ± 0.12
-(15-25)	7.0 ± 0.11
(-15)-(+15)	7.2 ± 0.13
15-25	7.4 ± 0.07
25-35	8.1 ± 0.11
35-45	8.6 ± 0.15
45-50	8.9 ± 0.18
Total	7.6 ± 0.08

Table A.7: The table shows the final variances of the ALPHA distribution for different zenith angle bins. The variances increase with increasing zenith angle.

way it can be made sure that most of the signal events are inside the cut limit and that the resulting measured flux no longer depends on the zenith angle.

The spectrum before and after applying this pointing correction remains the same which can be taken as a proof that no artificial effect has been introduced by optimizing on fluctuations.

A.7 Differential flux spectrum calculation

The differential flux is one of the main physical quantities that we are interested in. It is a measurement which can be compared with the theory and that tells us something about the mechanism that produces jets and high energy gammas inside the jet. The calculation of the differential flux is a rather complicated procedure that involves several steps.

1. The energy of the observed events is reconstructed. Then the ***excess events*** N^i_{Excess} for different energy bins are determined by estimating the background in each energy bin (this results in the ***excess event distribution***). The reconstructed energy can derivate, both systematically and statistically, from the true one.

2. Therefore the spectrum calculation requires the ***unfolding*** of these effects (of the excess event distribution).

3. Finally, the ***differential flux*** is obtained by dividing each bin by the total observation time T_{obs}, the bin width ΔE_{Bin} and its effective collection area. The observation time has to be taken separately for each zenith angle bin.

$$\frac{dF^i}{dE} = \frac{\Delta N^i_{Excess,unfolded}}{\Delta E^i_{Bin} \sum_{\theta_i} T^{\theta_i}_{obs} A^i_{eff}(\theta_i)} \tag{A.95}$$

A.7.1 Determination of the energy excess event distribution

The energy excess event distribution is the ***first step*** to the differential flux spectrum (dF/dE). The events are sorted in energy bins according to their estimated energy. The binning has been chosen to be of the size of the average energy resolution (~24 %). In ***logarithmic scale a constant binning*** (for simplicity) was introduced. The resolution is to first order constant in this frame (See Fig. A.38).

For each energy bin, a histogram for the ***ALPHA*** distribution was filled and the corresponding background was estimated by performing ***a fit*** as was done earlier.

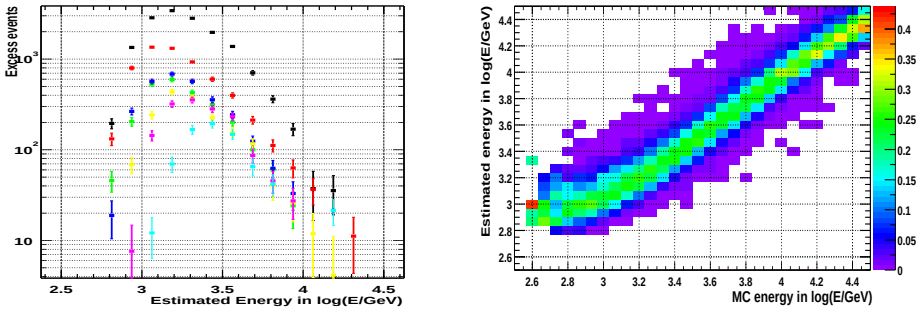


Figure A.38: a) This figure shows the ***excess event distribution*** of the total ***Mkn 421*** data set for all ***six*** zenith angle bins in different colors (red: 0°-16°, blue: 16°-21°, green: 21°-28°, yellow: 28°-33°, violet: 33°-40°, light blue: 40°-50°, black: all together). The figure illustrates how the threshold ***changes*** with zenith angle. Due to ***different*** observation times the total amount of excess events is different for different zenith angles. The black points are the ***sum*** of the data of all the zenith angles. b) The estimated energy and the MC energy are plotted against each other and normalized to one (Probability distribution of the estimated energy). It demonstrates that in first order the resolution ***is constant in logarithmic scale***. The curved shape of the distribution for energies below 1 TeV is corrected later on by the unfolding procedure.

$$N^i_{Excess,Data} = N^i_{Alpha<15\,deg} - N^i_{Backgound<15\,deg} \tag{A.96}$$

The ALPHA distributions and their background fit are listed in Appendix B. The systematic error introduced by the background estimation is believed to be smaller than 5%-10%. The effect of all possible systematic errors, which can affect the spectrum, will be discussed in detail in the conclusion. As will be shown, the spectral shape remains ***virtually unchanged*** by artificial changes in the amount of background of +-10%.

The energy distribution of the excess events can be seen in Fig. A.38. The different zenith angles are represented by different colors (red: 0°-16°, blue: 16°-21°, green: 21°-28°, yellow: 28°-33°, violet: 33°-40°, light blue: 40°-50°, black: all together). As expected the threshold increases with increasing zenith angle.

A.7.2 Unfolding the spectrum

The spectrum of a real source should be ***unfolded*** with the distribution of the estimated energy. This is important not only to correct for ***nonlinearities in the energy estimation***. It also corrects for ***spill over*** from each energy bin into the neighbor bins on the left and right side and a possible cutoff would change its position.

The energy estimation allows us to determine the transfer function of the telescope as the simulation describes the total system. As in Fig. A.34, we can fill a matrix $\mathbf{M}$ which yields us for each (real, simulated) energy bin the ***probability distribution*** (normalized to one) of the estimated energy. Naturally, the binning of $\mathbf{M}$ and the binning of the excess event distribution (i.e. data distribution D_k) must be the same. Knowing $\mathbf{M}$ we can can calculate the spectrum of the ***folded*** spectrum $\vec{F}$ for a given ***unfolded*** spectrum $\vec{U}$:

$$\mathbf{M}\vec{U} = \vec{F} \tag{A.97}$$

The reverse of this procedure is called ***unfolding:***

$$\mathbf{M}^{-1}\vec{F} = \vec{U} \tag{A.98}$$

Even though this procedure is mathematically valid, it does not usually yield decent results. Since the spectrum vector $\overrightarrow{F}$ and M also has statistical fluctuations (it has been determined by Monte Carlo studies), the resulting vector tends to artificial ***oscillations*** between ***neighboring vector entries*** (in 'energy'-space). There are various methods available which attempt to suppress these completely unphysical high frequencies by means of regulators or low pass filters in the frequency space.

The method used here is a stepwise ***iterative method*** which is especially suitable for our situation. It is described in [Des95]. The procedure has been slightly modified to work efficiently for our case. The algorithm systematically ***adjusts*** the ***folded MC*** distribution F_k to the ***real data*** distribution D_k until the chisquare between the two distributions becomes ***minimal.*** The original ***unfolded MC*** distribution is then the distribution that we are interested in. The regulation here consists of ***interrupting*** the iteration before it can develop high frequency oscillations.

The exact procedure is as follows:

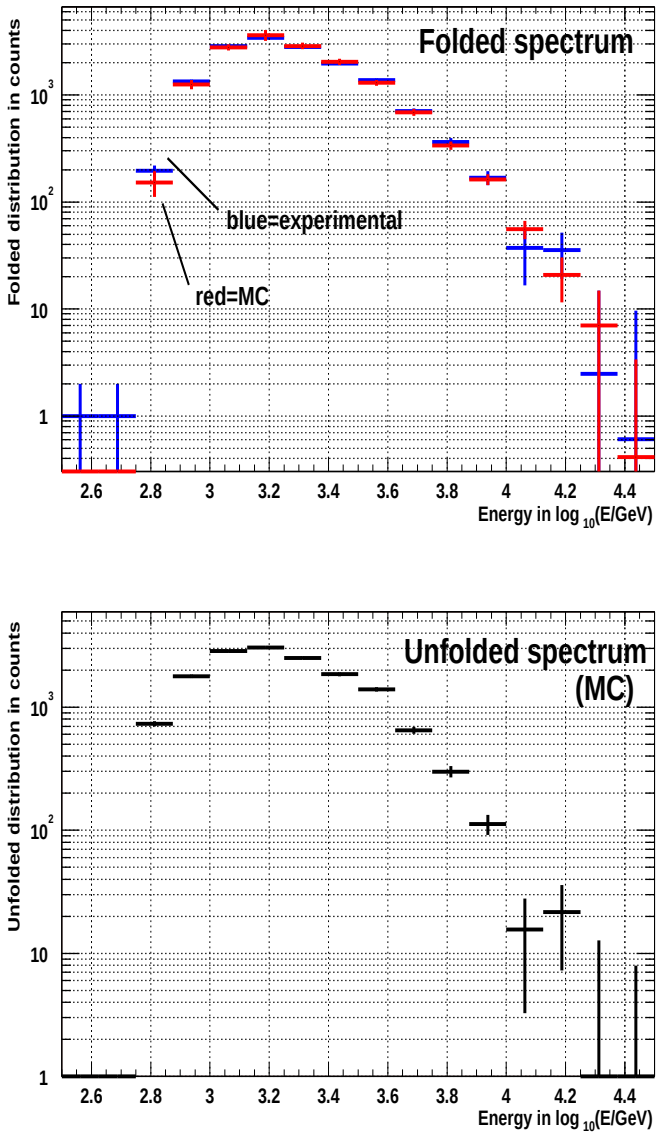
1. First a ***initial*** unfolded spectrum is ***estimated***. The closer the initial distribution to the final result, the better the convergence. In this case simply the ***folded data spectrum*** is taken (along with its errors) since it is already very similar to the expected unfolded result.

2. In an ***iterative loop*** the following steps l are executed for each bin i:

   (a) The ***folded spectrum*** is calculated from the unfolded distribution with (same as above)
   $$F_k^l = \sum_i \mathbf{M}_{ik} U_i^l \tag{A.99}$$

   (b) The ***ratio*** of the folded MC F_k and the real data distribution D_k is ***back-propagated*** into proportional factors λ_i
   $$\lambda_i^{l+1} = \sum_k \mathbf{M}_{ik} \left(\frac{F_k^l}{D_k^l} \right)^{\alpha_k} \tag{A.100}$$
   which are then applied to the ***unfolded data*** distribution
   $$U_i^{l+1} = N^{l+1} \lambda_i^{l+1} U_i^l \tag{A.101}$$
   in such a way that the ***folded*** MC distribution ***converges*** to the ***real*** data distribution. The row vectors (estimated energy distribution) of the matrix M_{ik} must be normalized in a manner such that for each index i the sum over all entries k give one. N^{l+1} is a normalization factor which ensures that the integral over the previous spectrum and after application of the proportional factors λ_i is preserved.

   (c) The ratio Equ. A.100 is potentiate with ***exponents*** $\alpha_k \in [0,1]$ which depend ***monotonically*** on the difference between the data distribution and the folded MC distribution, weighted by its error (χ^2 for one bin), in order to achieve ***smooth*** and equivalent ***convergence simultaneously*** along the ***entire*** spectrum.
   $$s_k = \frac{\left(F_k^l - D_k^l\right)^2}{\sigma^2_{F_k^l} + \sigma^2_{D_k^l}} \tag{A.102}$$
   $$\alpha_k = Sigmoid\left(s_k\right) \tag{A.103}$$
   $$Sigmoid\left(x\right) = \left(\frac{2}{1+e^{-x}} - 1 \right) \in [0,1] \tag{A.104}$$



(a) Unfolded spectrum

Figure A.39: The unfolding procedure is illustrated in the two plots. a) The experimental distribution (blue) is stepwise approximated by a folded MC distribution (red). b) The unfolded MC distribution is shown in the plot below. a) The folded spectrum changes its shape after unfolding especially in the energy region below 1 TeV and above 10 TeV ($log(E/GeV) = 4.0$) the cutoff appears stronger. Above $log(E/GeV) = 4.2$ the method becomes inaccurate due to very low statistics.

The sigmoid-like function has been introduced to assure that the exponents α_k are within the interval $[0, 1]$. Excessively large α_k values accelerate the convergence up to a point where the oscillations can no longer be controlled.

(d) The errors on the unfolded spectrum are calculated by ***Gaussian error propagation*** from the involved unfolded data bins and from the error on the transfer matrix M_{ik}.

(e) The χ^2 of the difference between the data spectrum D and the folded MC F is computed for each step

$$\chi^2 = \sum_k s_k \tag{A.105}$$

in each step. When the χ^2 starts to ***increase*** the iteration procedure terminates.

3. The procedure converges rapidly within ***five*** iterations. However, after a maximum of ***eight*** iterations the procedure is terminated to ensure that the high frequency oscillations do not develop.

Conclusion

The procedure begins with the excess event distribution obtained by using the energy estimation described above. The iteration process corrects the spectrum for ***nonlinearities*** and biases that were introduced by the energy estimation. In particular, it corrects for ***spill over*** effects from lower energy bins into higher energy bins. Without application of this unfolding procedure, the resulting spectrum would be flatter than the true one and a possible cutoff would be measured at a wrong position.

Because of statistical effects the unfolding procedure can ***only approximate*** the true spectral shape. This has been ensured by using the exponents α_k which are obtained by weighting with the variances (i.e. statistical fluctuations) in each bin.

The resulting distribution fits better to the effective areas that were obtained by MC studies as experience showed. Fig. A.39 demonstrates the result. The upper plot shows the original data distribution and the folded MC distribution. The lower plot shows the unfolded spectrum. Above 20 TeV virtually no significant signal (too low statistics) is seen so that this part of the distribution is discarded.

A.7.3 The spectrum, fitting and reverse check of the result

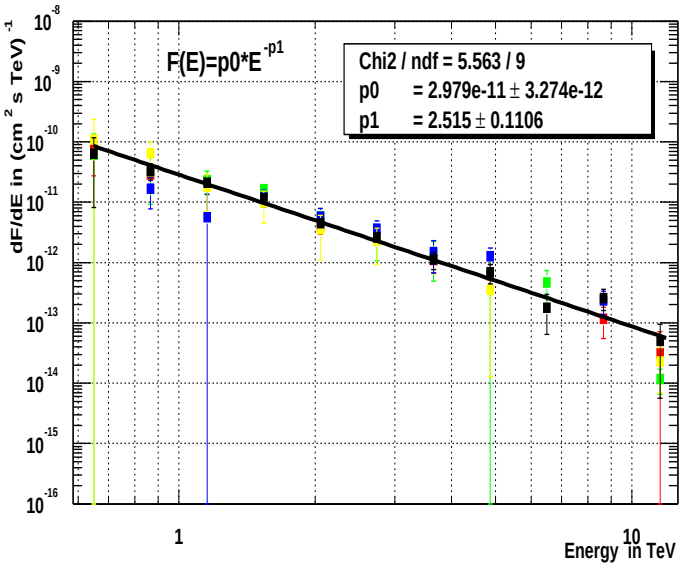
The data has been sorted into ***several zenith angle bins*** because the effective collection areas and the telescope threshold vary with zenith angle. The ***observation times*** for each zenith angle bin were carefully determined. In addition to the spectra for each zenith angle, ***the total spectrum*** was calculated by computing the effective areas and the normalized inversion matrix $\mathbf{M}_{ik}$ for the given zenith angle distribution as found in the dataset.

Fig. A.40 shows in the upper plot the spectrum of the crab nebula and in lower plot its excess event distribution which was used to obtain the spectrum. The spectrum (upper plot) has been fitted with a normal power law ansatz:

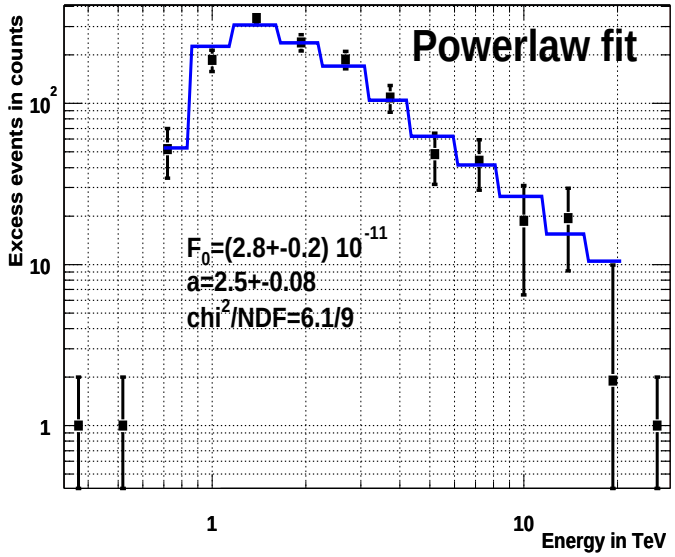
$$\frac{dF_{PowerLaw}}{dE} = F_0 E^{-\alpha} \tag{A.106}$$

with a differential spectral index α and a flux constant F_0. The ***fit results*** here are $F_0 = (3.0 \pm 0.3)10^{-11} TeV^{-1} cm^{-2} s^{-1}$, $\alpha = 2.5 \pm 0.1$ with a χ^2/NDF=5.5/9.

Since unfolding processes in general are ***not completely satisfactory*** and mathematically stable one might wish to ***cross check*** the obtained results. Therefore also the reverse process was also examined. Out of a given differential flux the expected number of excess events per energy bin $\left[E^k_{left}, E^k_{right}\right]$ can be computed:



(a) Crab spectrum



(b) Excess event distribution

Figure A.40: a) The upper plot shows the unfolded spectrum of the Crab nebula. The fit results here are $F_0 = (3.0 \pm 0.3)\ 10^{-11} TeV^{-1} cm^{-2} s^{-1}$, $\alpha = 2.5 \pm 0.1$ with a χ^2/NDF=5.5/9.
b) The plot shows the experimental excess event distribution of the Crab nebula. The theoretically expected excess event distribution for a power law spectrum (blue curve) has been fitted to the original excess event distribution (data points). Mathematically, this is more satisfactory because, as opposed to an unfolding process, is involved a folding process in this case and the expected errors are smaller. The fit results are: $F_0 = (2.8 \pm 0.2)\ 10^{-11} TeV^{-1} cm^{-2} s^{-1}$, $\alpha = 2.5 \pm 0.08$ with a χ^2/NDF=6.1/9. The values are consistent with the result from the upper plot. The fit error are slightly smaller due to the unfolding procedure applied in the upper plot. The Crab nebula is usually used as a standard candle and as a cross check to compare analysis results with other γ-ray experiments.

$$N^{Expec}_{Excess,i}\left(F0, \alpha\, E_{cut}, \beta\right) = \tag{A.107}$$
$$T_{obs} A_{eff,i} \sum_k \mathbf{M}_{ik} \int_{E^k_{left}}^{E^k_{right}} \frac{dF\left(F0, \alpha\, E_{cut}, \beta\right)}{dE} dE$$

The reverse process does not need an ***unfolding*** algorithm of the data spectrum, but rather a ***folding*** algorithm of an assumed unfolded spectrum. Mathematically, this is much more satisfactory, reliable and precise. A folded power law function has been fitted to the original excess event spectrum and can be seen in Fig. A.40 (lower plot). The ***fit results*** are: $F_0 = (2.8 \pm 0.2) 10^{-11} TeV^{-1} cm^{-2} s^{-1}$, $\alpha = 2.5 \pm 0.08$ with a χ^2/NDF=6.1/9. Both methods give practically the same fitting results and are therefore ***consistent.*** The fit to the original folded excess event distribution gives more precise results and ***smaller fitting errors.***

A.7.4 Discussion of systematic errors and the reliability of the obtained spectrum

The results of the analysis depend on several potential systematic errors which are discussed here.

1. The ***first*** error comes from the ***calibration of the absolute energy scale***. The pixels in the camera were calibrated by the so-called ***excess noise factor method*** which was already described above. Due to several unknown parameters and uncertainties, a miss-calibration of the absolute energy scale of up to 15 % is easily possible.
   A forced translation of the energy scale by a small amount (+- 10 %) changes the position of the threshold of the telescope relative to the threshold in the effective areas which consequently changes the shape of the spectrum in the region of the left slope of the excess event distribution. The resulting spectrum is no longer linear and results in a curved spectrum (in the threshold region), in either an upward or downward direction.
   However the total shape of the calculated spectrum is ***not very sensitive to a change of 10 %.*** The spectral index of the differential flux changes slightly. The spectrum of the crab nebula shows a clear ***power law*** spectrum as it should without any distortion of the power law close to the threshold. For this reason it is believed that the calibration of the absolute energy scale is ***not worse*** than 10 %.
   The change in the slope by artificially changing the energy scale by +-10% is only about $\sigma_{sys}(\alpha) = \pm 0.1$ for a spectral index of approximately $\alpha = 2.5$. Due to the uncertainty in absolute energy scale a possible cutoff position (Mkn 421) cannot be determined ***more precisely*** than approximately $\sigma_{sys}(E_{cut}) = \pm 1\, TeV$.

2. The ***second*** systematic error originates in the ***estimation*** algorithm for the ***background.*** The estimated miss-determination of the background is believed to be ***less*** than 10%. Surprisingly, a forced change of $\pm 10\%$ in the estimated background did not produce any ***significant change*** in the steepness of the spectrum. But for a forced overestimated/underestimated background the ***cutoff position*** (for Mkn 421) moves down/up by approximately $\sigma_{sys}(E_{cut}) = \pm 1\, TeV$. However, comparisons with the flux of the HEGRA CT-System and correlation studies with the RXTE/ASM satellite show that the origin of this analysis coincides with the ***origins*** of the other instruments. Therefore it is believed that that the background estimation is ***better*** or equal than 5%.

3. The ***third*** systematic error originates from the calculation of the effective areas which are calculated by applying the same ***selection cut*** to the data and to the MC.

This assumes that the MC describes showers and their imaging in a reasonable way. Systematic errors on the absolute flux due to slightly wrong effective areas are estimated to be approximately ***10 %*** or $\sigma_{sys}(F_0) = \pm 0.1\, F_0$.

A.8 Calculation of lightcurve and hardness ratio

The lightcurve is another important measurement which provides us with valuable information about the time structure of the flares of Mkn 421. It gives us hints about the size of this object and its emission regions and about the mechanisms that produce high energy gammas (SSC).

A.8.1 The mathematical background of integrated flux measurements

The ***lightcurve*** (Fig. A.41) consists of integrated flux measurements for small time bins:

$$F_{[E_{min},E_{max}]}(t) = \frac{dN}{dA\,dt} = \int_{E_{min}}^{E_{max}} \frac{dF(t)}{dE} dE \tag{A.108}$$

The so-called ***hardness ratio*** is the ratio of integrated fluxes of two energy intervals, an upper energy interval and a lower energy interval

$$R(t) = \frac{F_{\left[E_{min}^{upper},E_{max}^{upper}\right]}(t)}{F_{\left[E_{min}^{lower},E_{max}^{lower}\right]}(t)} \tag{A.109}$$

It is used to detect changes in the steepness of the spectrum, which might change with flare intensity.

For the lightcurve, the data is binned with the ***granularity a single run***. For each run bin, the duration T_{obs}, the average zenith angle θ and the center noon-MJD t are found. ***Three*** different ***energy intervals*** were treated simultaneously.

1. For the ***normal lightcurve***, an interval $[1.0,\, 20]$ TeV was taken. A hard cut at $E_{est} > 1.0$ TeV was applied in order to leave out the region that is too close to the energy threshold of the telescope so that the systematic error in the lightcurve could be reduced. This region is very sensitive to the zenith angle because of strongly varying collection areas there.

2. The other ***two energy intervals*** are needed for the ***hardness ratio*** calculation and are kept variable (this will be discussed later).

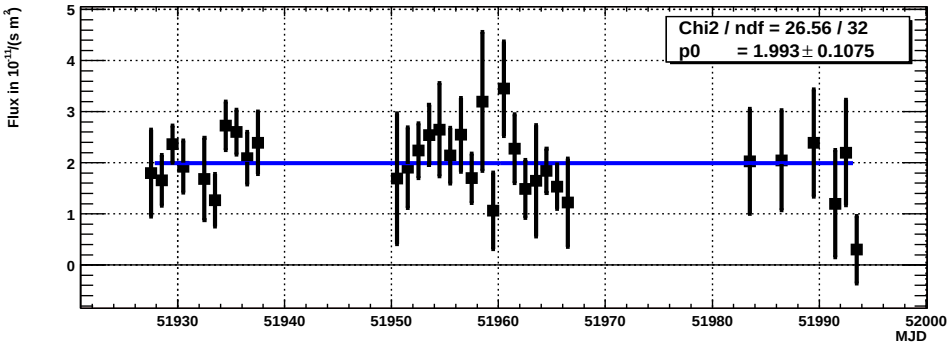
The ***excess-rate*** $Rate_{Excess}\,(t)$ and the background-rate $Rate_{Off}\,(t)$ are

$$Rate_{Excess}\,(t) = \frac{N_{On}\,(t) - N_{Off}\,(t)}{T_{obs}} \tag{A.110}$$

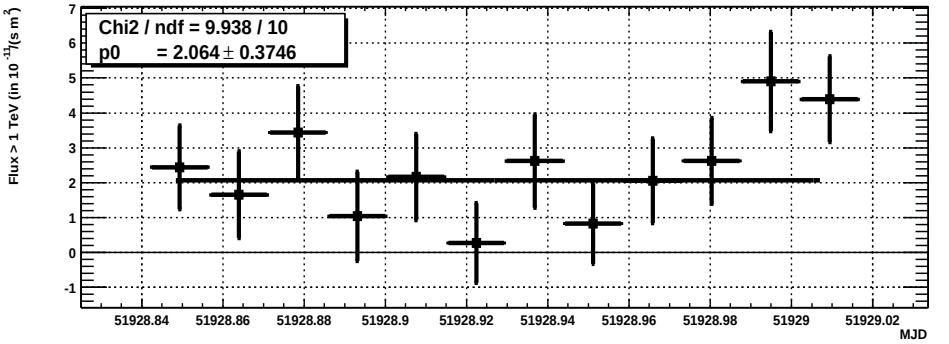
$$Rate_{Off}\,(t) = \frac{N_{Off}\,(t)}{T_{obs}} \tag{A.111}$$

The ***'On'-events*** are the events for which ALPHA is ***smaller than 18***°. The cut on ALPHA was chosen to be ***very generous*** in order to ensure that most signal events were retained so as to ***reduce*** the systematic error on zenith angle dependence (as was already mentioned above). The background events between 0° and 18° are ***estimated*** by fitting the ALPHA distribution for each run between 20° and 80° ***only*** with a polynomial. Due to low statistics, the ALPHA peak was not fitted with a Gaussian in this case.

The background rate is ***useful*** to detect ***corrupted*** runs (from bad whether or other problems). The background rate is slightly dependent on the zenith angle and decreases



(a)



(b)

Figure A.41: a) The light curve of the crab nebula from February to April 2001 averaged over one day bins and b) an example light curve of the crab nebula of one night with run flux points. The crab nebula emits a constant flux. According to the χ^2 value, the fluctuations are in the order of magnitude which would be expected from the errors on the flux points. This tells us that the error calculation of the lightcurve is consistent with the observed fluctuations.

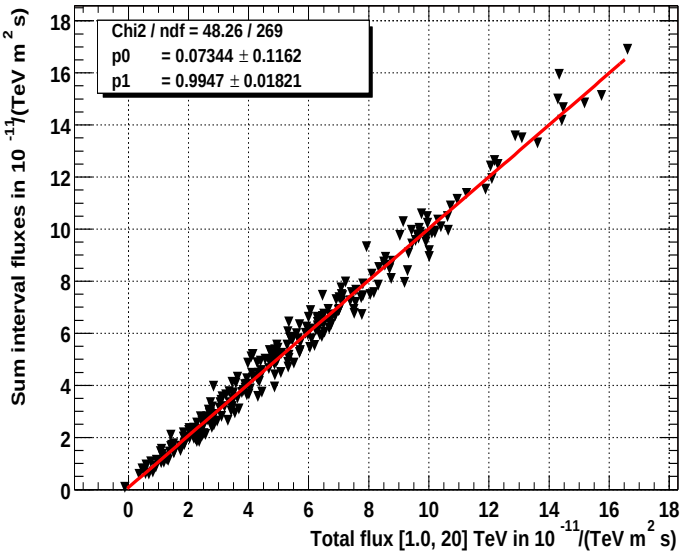


Figure A.42: ***Consistency check*** of ***lower energy*** interval flux ([1.0, 2.0] TeV) and ***higher energy*** interval flux ([2.0, 20] TeV) calculation (for the Mkn 421 data set). The ***sum of both*** fluxes is plotted against the ***total flux*** (energy interval [1.0, 20.0] TeV). The background estimation and also the effective area calculation for the curved spectrum is done separately for the different energy intervals. The plot proves ***consistency*** of the sum and the total.

with increasing zenith angle. All the runs with background rates of ***less than 10 events/hour*** (usually the rates are in the order of 40 events/hour) are ***rejected*** in a quality selection cut.

Hence, the ***integrated flux*** for a given energy interval is defined as the excess rate for this energy interval divided by its ***average effective area*** for the corresponding zenith angle:

$$F_{[E_{min},E_{max}]}(t) = \frac{Rate_{Excess,[Emin,Emax]}\,(t)}{A_{eff,int}\,(\theta)} \tag{A.112}$$

In order to obtain the average effective area the ***differential fluxes*** from the spectrum calculations are used because these numbers are ***very dependent*** on the ***shape*** of the spectrum. The effective area is ***interpolated*** for the zenith angle needed. The average zenith angle changes from run to run. Figure A.42 shows in a cross check the consistency of the flux calculation. The sum of the lower energy [1.0, 2.0] TeV and higher energy [2.0, 20] TeV interval fluxes are plotted against the total flux (energy interval from [1.0, 20] TeV).

The average effective area is a ***weighted average*** of the flux summed over all energy bins inside the given interval [WitCom].

$$A_{eff,int}\,(\theta) = \frac{\sum_{E^i\in[E_{min},E_{max}]} \frac{dF}{dE}\left(E^i\right)\cdot(E_{max}-E_{min})\cdot A_{eff}\left(E^i,\theta\right)}{\sum_{E^i\in[E_{min},E_{max}]} F^i}$$

The events are sorted into the ***energy intervals*** by using the estimated energy. For this calculation of the lightcurve ***no unfolding*** was performed due to ***very low statistics***. Since rather large energy intervals are integrated, this should not introduce excessively large systematic errors. The only potential systematic error originate from the region of the energy threshold that is ***largely excluded*** (with the cut at 1 TeV).

The error in the flux is derived from Gaussian error propagation and its nature is purely statistical.

$$\sigma^2_{Rate} = \frac{N_{On} + N_{Off}}{T^2_{obs}} \quad \text{(A.113)}$$

$$\sigma^2_F = \frac{\sigma^2_{Rate} + F^2\sigma^2_{Aeff}}{A^2_{eff,int}} \quad \text{(A.114)}$$

A.8.2 Discussion about systematic errors and the reliability of the measured lightcurve

Two potential systematic errors can influence the lightcurve measurement.

1. The first one results from the ***impossibility*** of performing an unfolding procedure on the basis of run, due to a lack of available statistical information. This error can be significantly ***reduced*** a lot by performing a hard cut on the lower energy region.

2. The second one is related to the ***background estimation***. As explained in the section concerning the differential flux measurement, a maximum systematic error of 10% is assumed. However, comparing the flux results to the one of the HEGRA CT-system and also to correlation studies with the RXTE/ASM satellite (both in the next chapter) show that the ***origin*** of the flux calculation is in good agreement with the other two instruments. For this reason it is believed that the background estimation is better than or equal to 5%.

Appendix B

Analysis of the flares of Mkn 421 in 2001

In chapter 1 and A the necessary foundation was established to now discuss the analysis of the Mkn 421 data. This data was recorded between ***February 2001 and May 2001*** when Mkn 421 exhibited large flares. The total observation time amounts to 249 hours.

Two main characteristics will be examined. One is the lightcurve which contains ***timing information*** of the flares and time correlations to x-ray measurements. The second concerns the ***spectral properties*** of the flux. Calculated will be the total average spectrum and the spectrum in different flare states including hardness ratio changes.

B.1 Light curve and timing of the flares of Mkn 421

In this section, the measurement of the lightcurve using the ***CT1 telescope*** and other instruments like the ***x-ray satellites RXTE*** and ***ASCA*** and the ***CT-system*** telescopes are described and compared. In addition, fundamental timing properties and time correlations are discussed here.

B.1.1 The light curve of Mkn 421 as measured by CT1

The algorithm of the lightcurve calculation was explained in detail in chapter A. The data is ***binned in runs*** which are ***typically of 20 minutes duration***. Smaller binning was disregarded because the statistics would be unacceptably low. The lightcurve of the whole period is shown in Fig. B.2 and the complete night by night set can be seen in Appendix B. Some runs have high fluxes of up to $(10-15)\ 10^{-11}(TeV\ m^2\ s)^{-1}$. The lightcurve exhibits five main observation periods which are interrupted by moon periods during which observations was only restricted possible. In three of these five periods Mkn 421 showed high fluxes. Moon observation data is not included in this plot and was not analyzed in this document. A lightcurve with daily averages was already shown in Fig. A.41.

Estimation of rise and fall times

Measurements taking during twelve days and containing significant flares were selected and are shown in Figures B.3, B.4 and B.5. There is some evidence that some of the flares are faster than the resolution of the run bins. It is difficult to estimate the rise time in some cases. To obtain numbers for the doubling rise time, the halving fall time and the average FWHM duration of a typical flare, ***three*** different approaches were used.

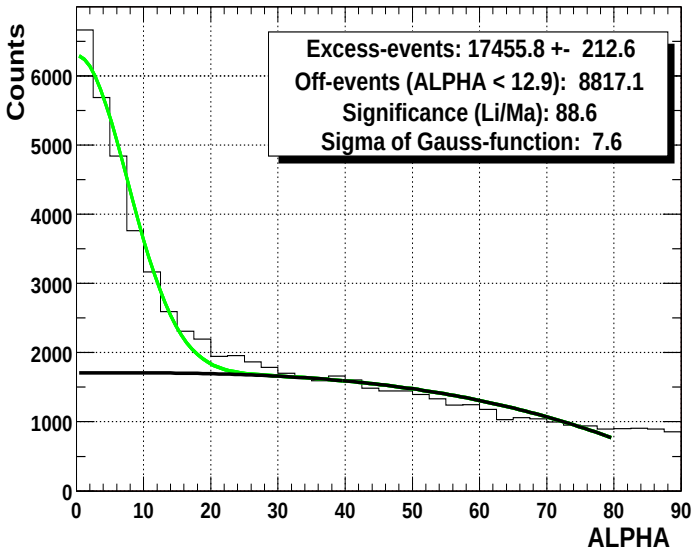


Figure B.1: This figure shows the total excess of the Mkn 421 dataset which amounts to 17400 +- 210 signal events over approximately 8800 background events and which was recorded between February 2001 and May 2001 with a total observation time of 249 hours.

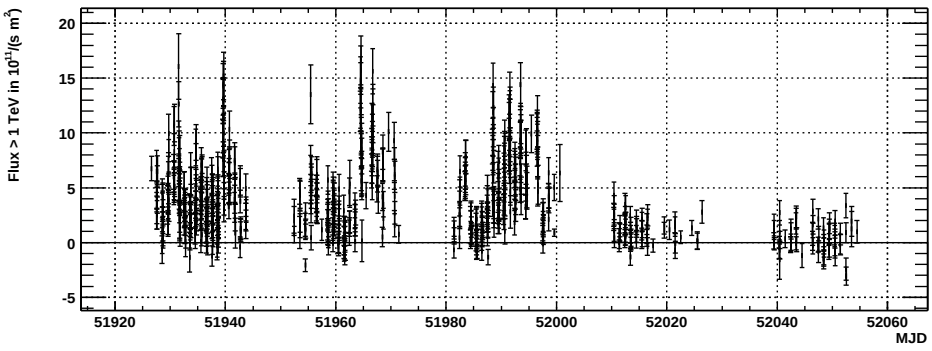


Figure B.2: The light curve for the complete period between February 2001 and May 2001 is shown. Each point corresponds to a single run which is typically of approximately 20 minutes duration.

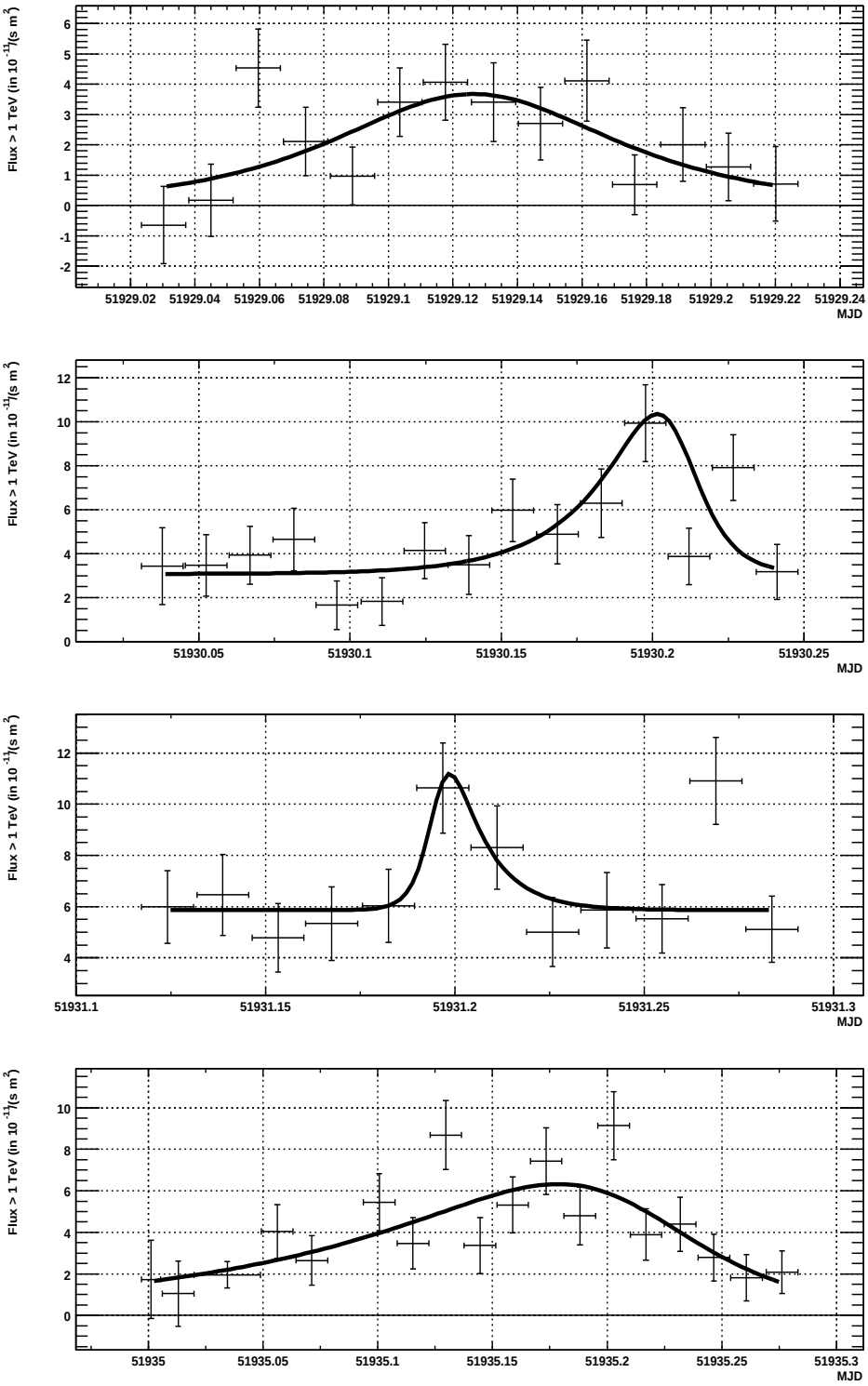


Figure B.3: Selected flux measurements from the nights 51928, 51929, 51930 and 51934. The flares have been fitted with a simple flare model described below.

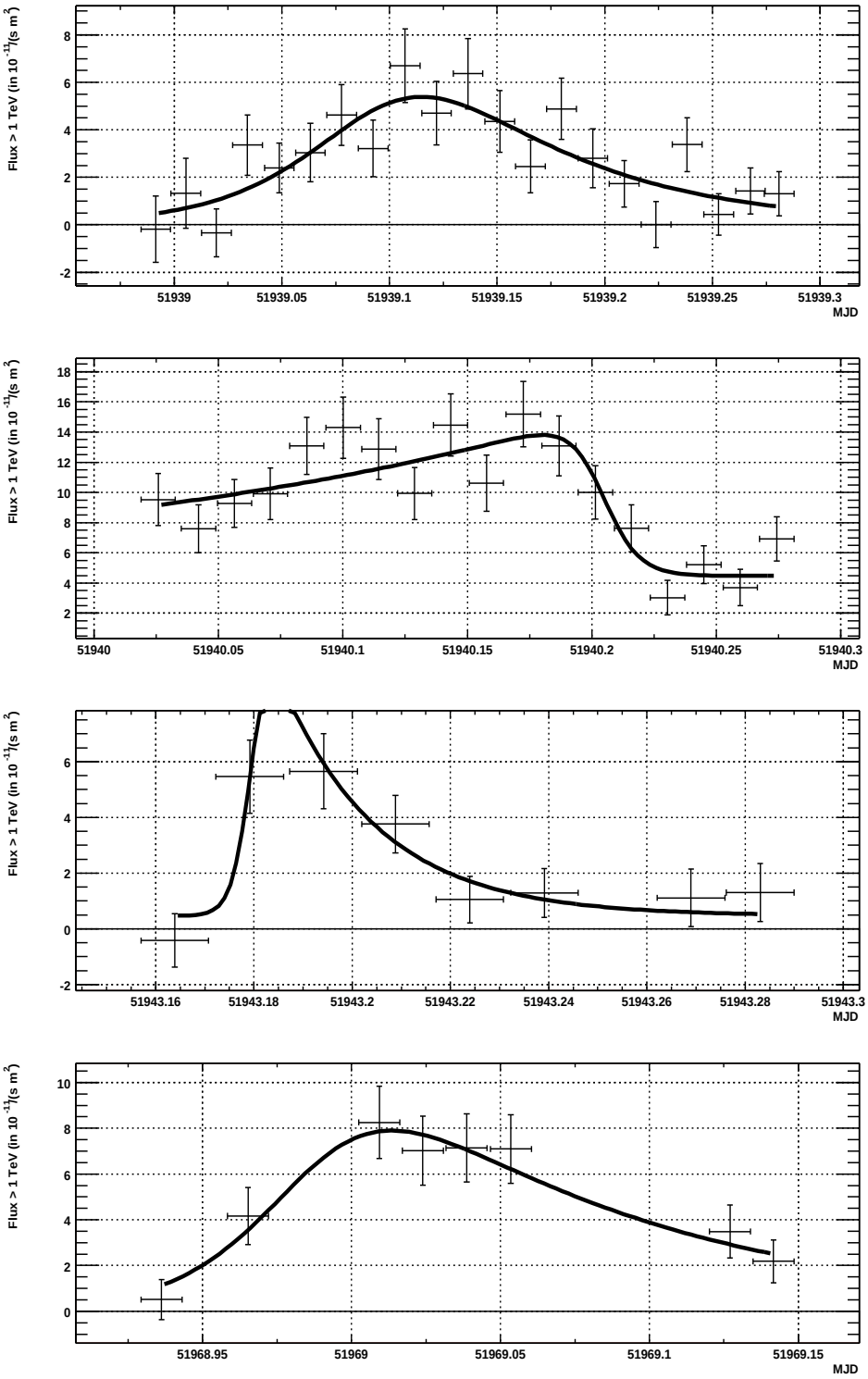


Figure B.4: Selected flux measurements from the nights 51938, 51939, 51942 and 51968. The flares have been fitted with a simple flare model described below.

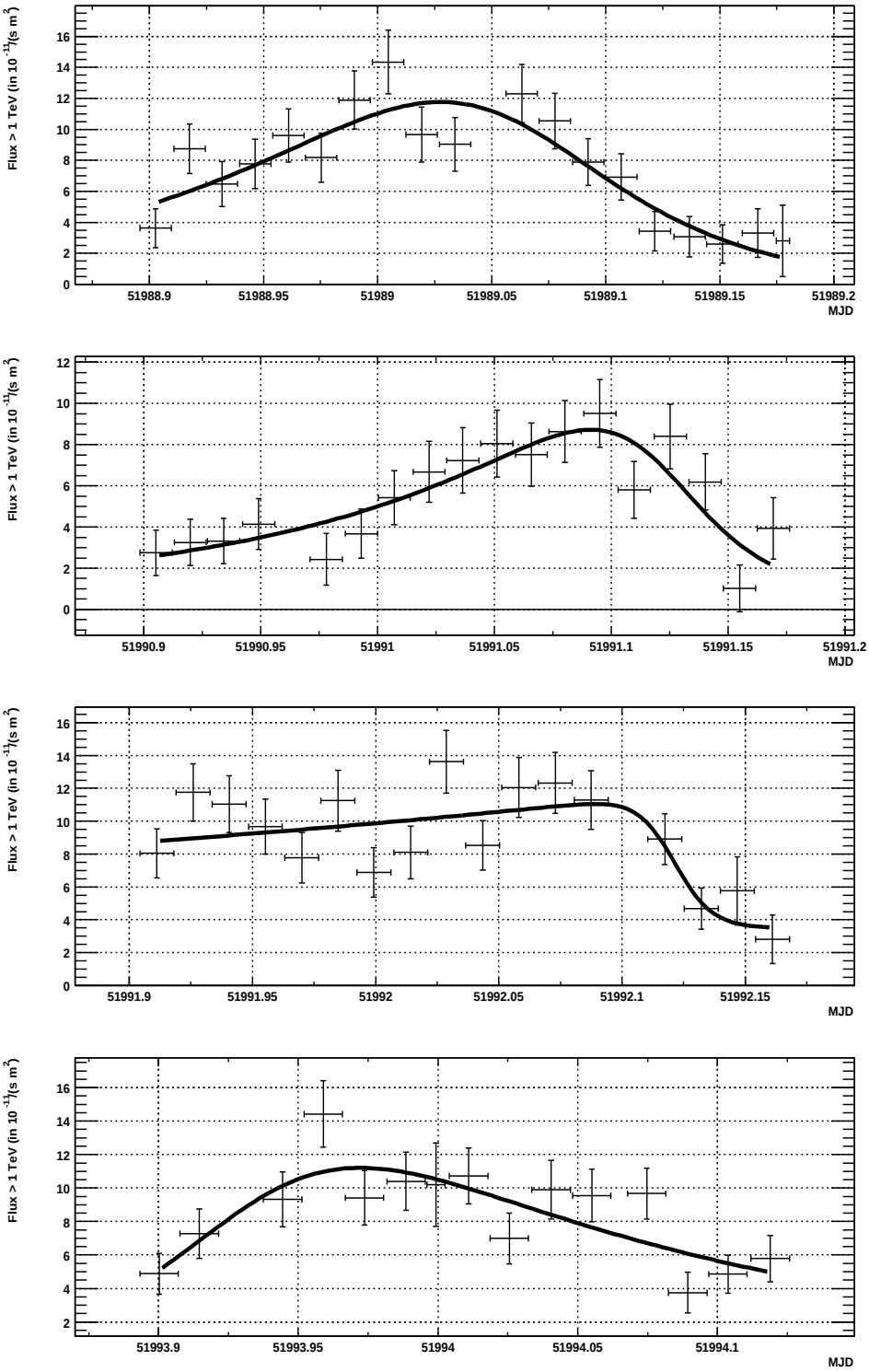


Figure B.5: Selected flux measurements of the nights 51988, 51990, 51991 and 51993. The flares have been fitted with a simple flare model described below.

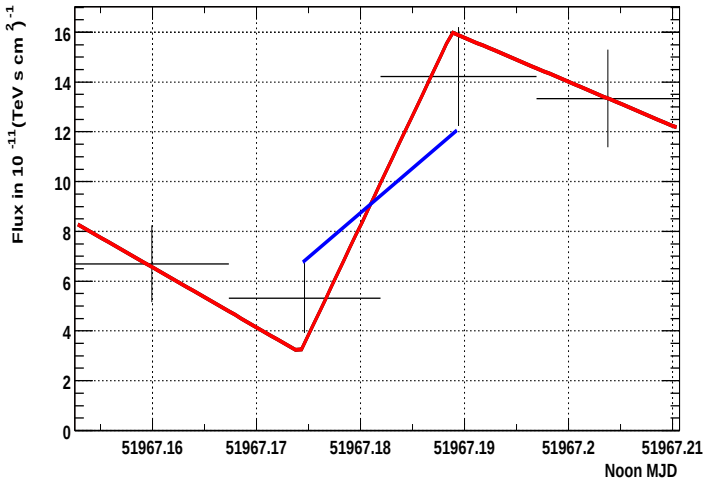


Figure B.6: Fit to obtain the doubling time. The flare taken occurred in the night 51967. It shows a doubling time of 8.2±7.0 minutes. Even the worst case scenario (blue line) still has a doubling time of only 23 minutes.

1. For a simple characterization, ***a flare model*** has been fitted to each night.

$$F(t) = a + \frac{b}{2^{\frac{(t-t_0)}{c}} + 2^{-\frac{(t-t_0)}{d}}} \tag{B.1}$$

   This function behaves like an exponential in the limits of minus and plus infinity. Rises and falls are supposed to have exponential character (corresponding to sudden outburst and exponential cooling). It has doubling rise and halving fall times c and d. This model assumes that several small fast flares ***pile up*** to a single very large flare. This means that the flare doesn't start from zero but rather from a ***constant background*** which is estimated by a. This fit is used to estimate the average duration of fast flares and the rise and fall times starting from a constant background.

2. For very fast flares ***four points*** are taken with the rise or fall in the middle (see Fig. B.6). Three segments of straight lines are fitted to these four points by ***integrating*** over the run period. In this manner we can approximate the slope of the middle segment by using the information from the points on the left and the right and obtain its error from the fit. The fit contains ***two free parameters*** which are the slopes of the first and second segments. The offset from zero on is given by the first and last point which lay on the first and last straight line. The doubling time is calculated here by assuming the a flare starts from ***zero*** and not starting from an offset as by fitting the flare model. In this way the rise time is calculated ***conservatively***.

3. Since the 4 point fit is not useful in all the cases, the doubling rise time between ***two neighbouring points*** is calculated by means of a ***simple line*** connection between two points. The error is obtained from the errors of the fluxes. Since this method is very sensitive to statistical fluctuations, the doubling and falling times are calculated by taking the ***worst case scenario***. The flux values used are the measured fluxes plus/minus their one sigma errors. (The ends of the error bars are connected, see Fig. B.6).

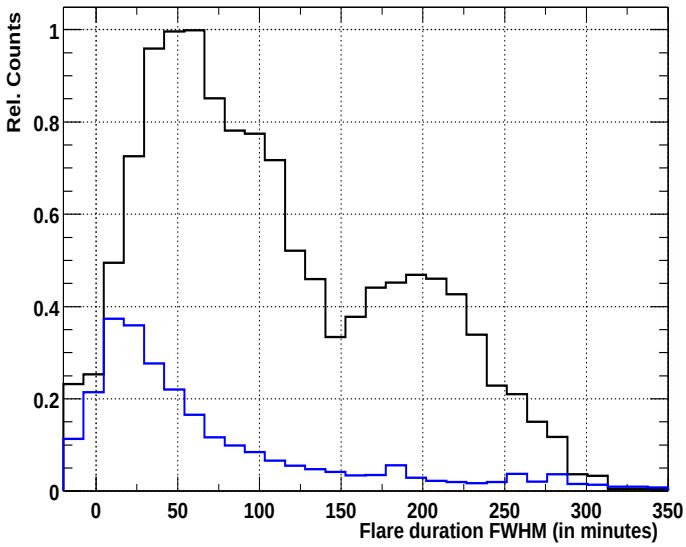


Figure B.7: FWHM duration of 21 large flares obtained by fitting the flare model above. The typical duration for fast flares ranges between ***one hour*** and ***three and a half hours***. The distribution of the duration is obtained by filling ***Gaussian distributions*** into a histogram. The mean and sigma of such an Gaussian distribution corresponds to the measured value and its errors. The ***underlying blue line*** is the ***noise background*** that was obtained by simulating a lightcurve with Gaussian random distributed data points around a mean of the daily average flux. Then the same flare model was fitted as to the real light curve. For each night ***100*** lightcurves were simulated. The blue line shows that for random lightcurves ***significantly less*** flares are found.

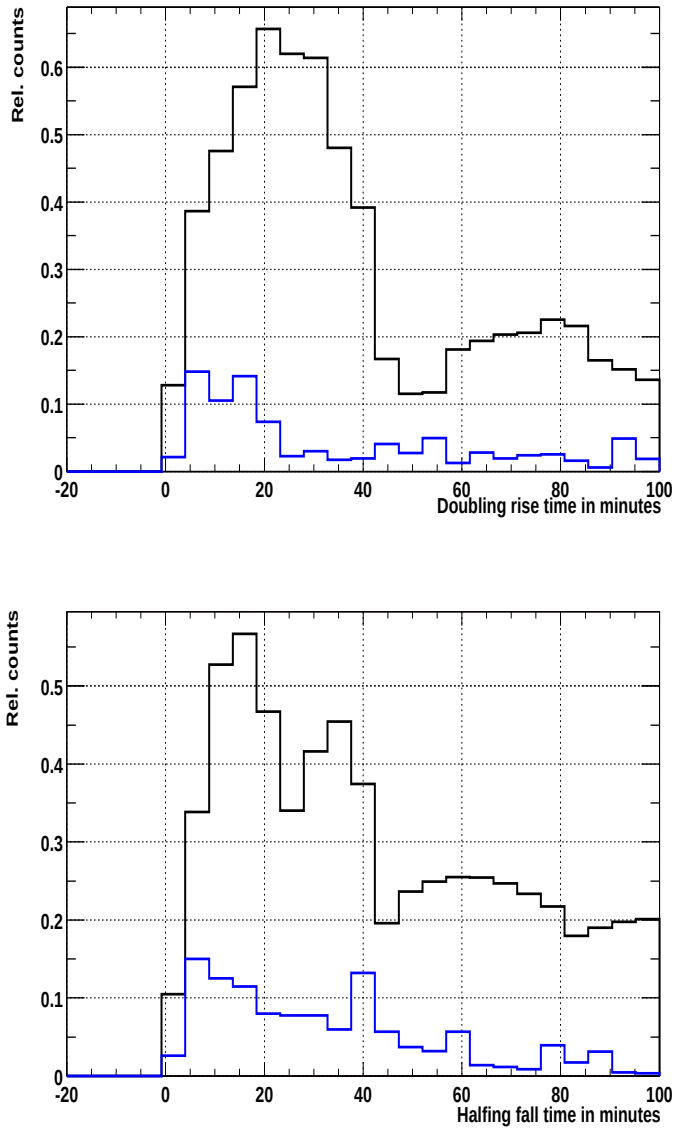


Figure B.8: These two plots show the parameter values c (doubling rise time) and d (halving falling time, see Equ. B.1) obtained from fitting 21 large flares. Typical time scales in this model range from upwards 25 minutes. The distributions are obtained by filling ***Gaussian distributions*** into a histogram. The mean and sigma of the Gaussian distribution corresponds to the measured value and its errors. The ***underlying blue line*** is the ***noise background*** that was obtained by simulating a lightcurve with Gaussian random distributed data points around a mean of the daily average flux. Then the ***same*** flare model was fitted as to the real light curve. For each night 100 lightcurves were simulated. The blue line shows that for random lightcurves ***significantly less*** flares are found.

Type	MJD	Doubling/halfing time (in minutes)
Fall	51967.05	11.7 +- 11.3
Fall	51990.06	14.1 +- 16.3
Fall	51991.12	12.8 +- 14.9
Rise	51934.18	13 +- 10.7
Rise	51956.06	15 +- 15.2
Rise	51967.17	8.24 +- 6.9

Table B.1: The results of the four point, three segment fit are presented here. The shortest fall time is 12.8 +- 14.9 minutes and the shortest rise time is 8.24 +- 6.9 minutes. Taking into account the large errors, there is evidence of flares that have rise and fall times of 15 to 25 minutes.

Fitting the flare model

The starting values for the fit are set as follows: a=average flux during night, a=8.0, c=30 minutes rise time, c=30 minutes fall time and t_0=the position of the highest flux in that night. In order to select nights with significant flares, first a straight line fit is performed. If the reduced chi-square is less then 1.5, $\chi^2/NDF < 1.5$, then the night is not used for the fit of Equ. B.1. The most interesting major flares are shown in Fig. B.3, B.4 and B.5 while the complete set can be seen in Appendix B.

Only fitting values with errors less than 100% are allowed (as a quality selection cut) for the reminder of the process. 21 fitted flares fulfilled this criteria. The plots in Fig. B.8 and Fig. B.7 are obtained by filling area-normalized ***Gaussian distributions*** into a histogram. The mean and sigma of the Gaussian distribution corresponds to the measured value and its error, respectively. In this way, it is possible to graphically illustrate the distribution for only few measurements. In order to estimate the ***significance*** of the obtained result, the same procedure was applied to lightcurves that were generated by a ***Gaussian random generator***. The mean of the Gaussian distribution is the average daily flux and the variance is the error of each data point. For each night 100 lightcurves were simulated and fitted by the same procedure like the real one. The obtained distribution shows that in random distributed lightcurves ***much less*** flares are found and thus the measured rise time distribution is significantly higher than the background.

The ***average duration*** of fast flares ranges from ***one hour*** to ***three and a half hours*** FWHM. The typical doubling rise and halving fall times in this model range from upwards of 25 minutes. A significant difference between rise and fall times is not observed.

Four point-three segment fit

The results of the four point, three segment (three straight lines) fit are presented here. A sample of rise and fall times faster than 15 minutes are listed in Tab. B.1. The flares can be found in the Fig. B.3, B.4 and B.5 by its MJD.

Taking into account the considerable errors on the measurements, there is evidence that there exist flares that have doubling and halving times of only 15 to 25 minutes.

Worst case scenario: Two point straight line connection

Since the four point fit with two free parameters introduces large errors and since not all fast flares have two neighboring points, the rise and fall time were also calculated from a straight line connection, but ***subtracting one sigma*** from each of the two points (the worst case scenario). Some of the points which are faster then 20 minutes are listed in Tab. B.2.

When considering the errors one should take into account that in this ***worst case scenario*** already one sigma had already been subtracted/added on each of the two points and one obtains still a rise time of 8.24 +- 7.0 minutes. Therefore it can be safely stated that

Type	Noon MJD	Doubling/halfing time in minutes
Rise	51932.01	8.31+- 4.9
Rise	51989.90	9.7 +- 6.6
Fall	51929.05	19.4 +- 13.8
Fall	51992.14	9.57 +- 7.5

Table B.2: In this table some fast rise and fall times are listed. They have been calculated out of a two point connection from which one sigma has been subtracted already (worst case scenario, see Fig. B.6). It can be stated that in a few cases the ***doubling rise and halving fall times*** seem to be ***equal or faster than 15 minutes***.

in a few cases the ***doubling rise and halving fall times*** are for sure ***equal or faster than 15 minutes***.

Conclusions from lightcurve, rise and fall time measurements

The lightcurve measurement made by the CT1 telescope shows a ***very fast variability*** of the source. There is evidence that the fastest ***doubling times*** and halving times are ***equal or faster than 15 to 25 minutes.*** The fit of a flare model to the lightcurve shows that the typical ***FWHM duration*** of the fast flares are ***between one hour and three and a half hours*** while in that model the ***typical*** doubling/halving times are approximately upwards of 25 minutes. A significant difference between rise and fall times has not been seen.

In addition, high fluxes of up to $(15 \pm 2) \cdot 10^{-11} TeV^{-1} s^{-1} m^{-2}$ can be observed sporadically during the four months flare period.

B.1.2 Comparison of the lightcurve to the one measured by the CT-system

To test ***redundancy***, the obtained flux was compared with the flux obtained by the ***CT-system***. This could only be done for time-coincident bins. As Mkn 421 is a very rapidly flaring source, if the flux is not measured by both the CT1 and the CT-System exactly at the same time then it cannot be expected that the measured fluxes precisely coincides. This results in a few points that don't fit perfectly in the trend. However, it can be checked whether both measured fluxes are correlated, whether the scale is the same and whether the origins coincide as well. Fig. B.9 shows the correlation of time coincident bins of quarter day size. The measurements are in good agreement. A straight line fit with $y = p0 + p1 * x$ to the data shows that the origins of both telescopes coincide ($p0 \sim 0$). ***The CT1 flux is approximately 13%$\pm$6% lower than that one determined by the CT-System.***

It should be mentioned that a completely different calibration procedure for CT1 and and CT system has been used. In this respect the difference of 13%$\pm$6% is surprisingly good.

B.1.3 The 2 keV-12 keV lightcurve measured by the All Sky Monitor of the RXTE satellite

Flares of February 2001 to April 2001 have also been detected in the ***x-ray domain at 2 keV to 12 keV*** by the All Sky Monitor (***ASM***) of the ***RXTE*** satellite. This allows ***multi-wavelength observations***. Some basic predictions of the SSC model can be tested. It claims that X-rays are produced by synchrotron radiation of said accelerated electrons while the GeV/TeV γ's are produced by inverse Compton upscattering of said x-ray photons.

Unfortunately the measurements of the All Sky monitor are only a few, short, 90 s measurements per day with very poor statistics and big errors. This does not allow precise

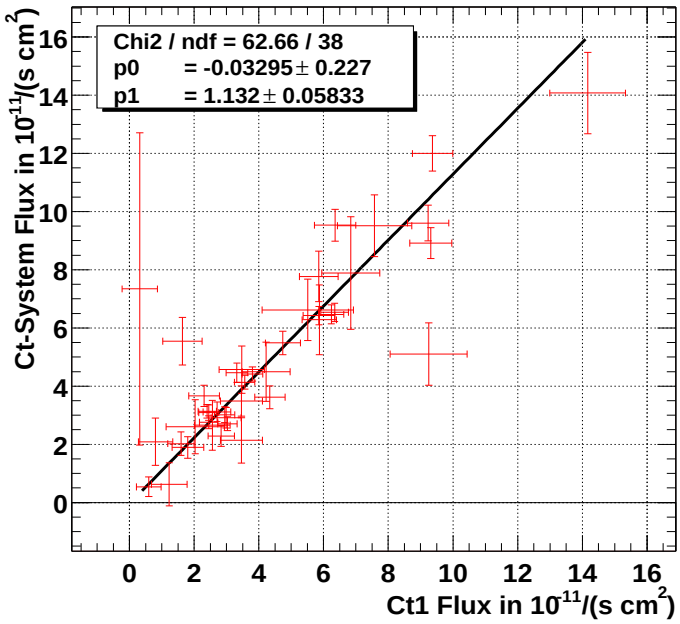


Figure B.9: Correlation of the flux above 1 TeV obtained by the CT-system when compared to the fluxes of CT1 for time-coincident bins of quarter day size. A straight line fit with $y = p0 + p1 * x$ shows that the ***origins of both telescopes coincide*** and that the CT1 flux is approximately ***13% smaller*** than the one of the CT-System. The points which are significantly away from the straight line fit are a result from not completely time coincident measurements of the two instruments.

measurements of the development of the flares. However the measurements can be averaged over bigger time bins. Figure B.10 shows (middle plot) the ASM lightcurve averaged over 12 hour bins and (above) a 24 hour sample of the day 51966. On the small time scale the data is highly noisy. Almost simultaneous points show large differences. If the data is averaged over a bigger time period acceptable results are obtained. In the bottom plot the corresponding CT1 half day bin lightcurve is displayed for comparison. In the next session the time correlation between the two instruments will be discussed in detail.

B.1.4 The discrete correlation function applied to TeV and X-Ray light curves

Looking at Fig. B.10, one may surmise that the fluxes of CT1 and RXTE are correlated with each other. Such a correlation provides very important information about the processes that create x-ray photons and γ-photons.

The empirical correlation function

The ***empirical correlation*** ρ of two vectors F^γ and F^x is generally defined as:

$$\rho = \frac{\sum_i \left(F_i^\gamma - \langle F^\gamma \rangle\right)\left(F_i^x - \langle F^x \rangle\right)}{\sqrt{\sum_i \left(F_i^\gamma - \langle F^\gamma \rangle\right)^2 \sum_i \left(F_i^x - \langle F^x \rangle\right)^2}} \tag{B.2}$$

In our case the situation is more complicated since the measurements are completely non-equidistant distributed in time. Both datasets have been re-binned into ***equidistant*** time bins of size ΔT, as described in Equ. E.35 (Appendix A). The correlation is measured for ***systematic time shifts*** Δt between the two datasets, the flux bins for γ-rays $F^\gamma\left(t_i\right)$ and for x-rays $F^x\left(t_i\right)$, to see if the light of the γ-flares arrives earlier or later than the light of the x-ray flares. Only the time coincident time bins are summed up, using the following formula:

$$\rho\left(\Delta t\right) = \frac{\sum_{t_k = t_i + \Delta t}\left(F^\gamma\left(t_k\right) - \langle F^\gamma \rangle\right)\left(F^x\left(t_i\right) - \langle F^x \rangle\right)}{\sqrt{\sum_{t_k = t_i + \Delta t}\left(F^\gamma\left(t_k\right) - \langle F^\gamma \rangle\right)^2 \sum_i \left(F^x\left(t_i\right) - \langle F^x \rangle\right)^2}} \tag{B.3}$$

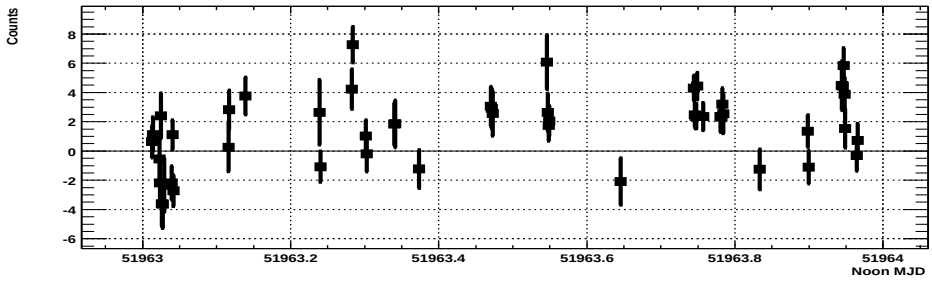
The finer the bin size is chosen, the fewer coincident time bins are found. The error increases because the fluctuation of smaller binned measurements are larger and because less time-coincident bins are found. The error calculation of the empirical correlation function is described in Appendix A.

Correlation of the data recorded by CT1 and ASM in 2001

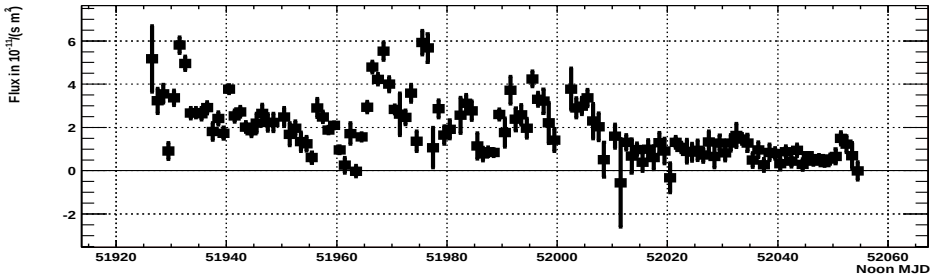
Fig. B.11 shows the correlation for time shifts between ± 40 days. In Fig. B.13 the RXTE flux and the CT1 flux are plotted against each other for time coincident bins and Fig. B.12 shows both lightcurves for comparison. For better visibility, only coincident time bins are included in the plot. Tab. B.3 shows the ***correlation for different bin sizes***. It ***increases*** with bin size.

Correlation of the data recorded by the ASCA x-ray satellite in 1998 and several Cherenkov-telescopes

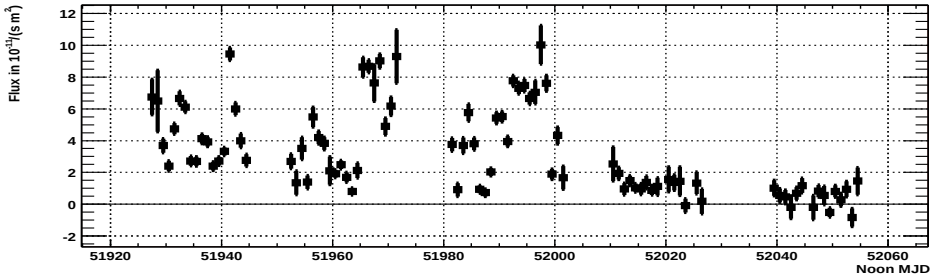
In the introduction chapter the measurement of the ***ASCA x-ray satellite*** and several Cherenkov telescopes (Wipple, HEGRA, CAT) between April 23, 1998 and May 1, 1998 were already presented (see Fig. B.14). The ASCA measurement is significantly more detailed and precise than that of ASM. Unfortunately, the TeV data is presently very sporadic and contains large error bars. However, the ***correlation*** between the x-ray flux and γ-flux was also analyzed.



(a) ASM (one night)

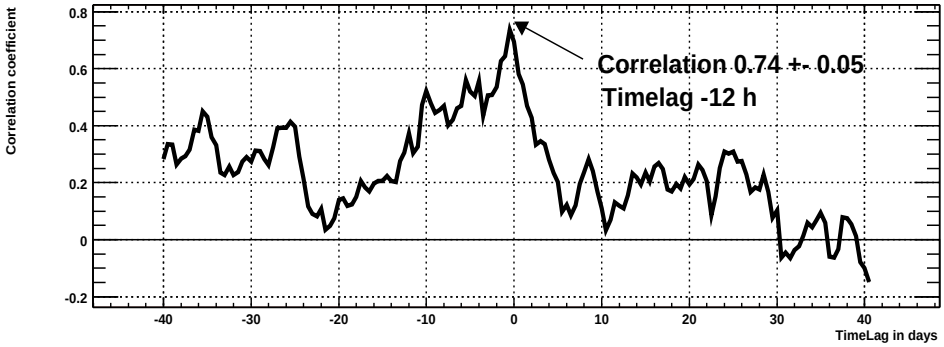


(b) ASM total period

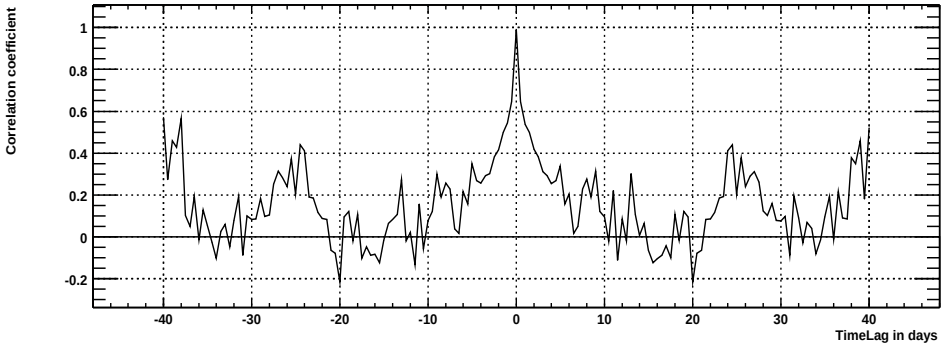


(c) CT1 total period

Figure B.10: The upper plot shows: a) A 24 hour sample of a ASM measurement for the day 51964. It can be seen that the ASM data is ***very noisy***. Almost simultaneous points show large differences. The middle plot shows: b)The light curve ***averaged over 12 hour bins***. The lower plot shows: c) The CT1 light curve for comparison. The ASM curve is continuous while the CT1 observation is interrupted by moon periods. From the upper two plots it is evident that while the ASM is very noisy when averaged over 12 hour bins it displays the ***trend*** of the flux. The flux outbursts seem to happen simultaneously for x-rays (ASM) and for γ-rays (CT1).



(a) ASM-CT1 correlation



(b) Autocorrelation of CT1

Figure B.11: a) The upper plot shows the correlation between x-ray data and the γ-ray flux, as measured with CT1. The ***maximum correlation*** 0.74 ± 0.06 is reached at a time lag of the x-rays of -12h. It has to be mentioned that in this plot the timelag is not significant. The time lag will be discussed later in detail. The bin size was 1.0 days. To obtain a smooth curve, the time lag between bins was sampled in steps of 1/3 days. b) The lower plot is the autocorrelation of the CT1 data. The data seems to contain periodicities, which also appear in the upper plot.

Bin size (days)	Correlation
5.0	0.91 ± 0.07
2.0	0.83 ± 0.06
1.0	0.74 ± 0.05
0.5	0.73 ± 0.05
0.2	0.65 ± 0.05

Table B.3: The table shows the correlation for different bin sizes. It ***increases*** with bin size. It cannot be clarified in this thesis whether this has a physical origin or if this originates from larger errors when using smaller bin sizes.

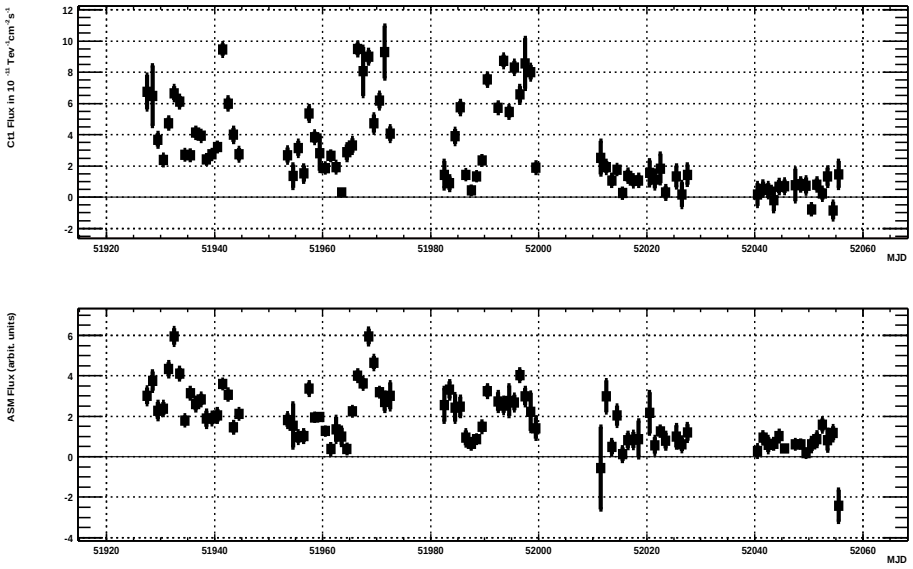


Figure B.12: a) The upper plot shows the CT1 flux and b) the lower plot shows the ASM flux. Note that, for better visibility only the time coincident bins of CT1 fluxes and ASM fluxes are plotted. The correlation is clearly visible now.

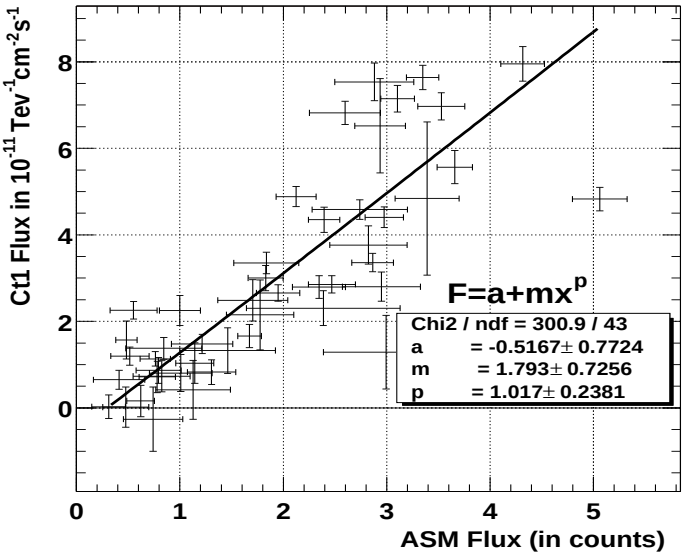
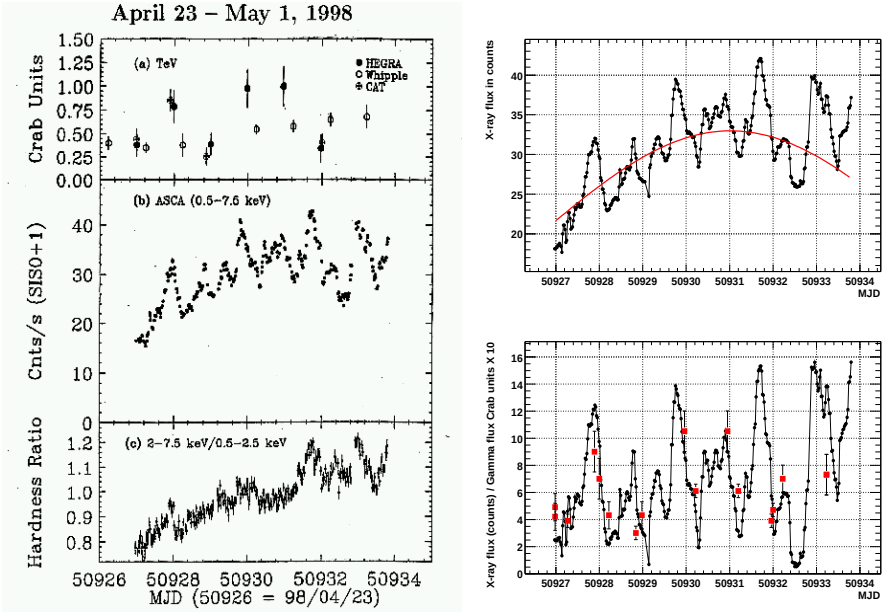


Figure B.13: In this plot the time coincident measurements are plotted against each other using a bin size of two days. The y-axis gives the CT1 flux in and x-axis the ASM x-ray flux in counts ($\propto$Flux). A function defined as $F = a + mx^b$ was fitted to the data. It reveals an ***almost linear*** dependence of the ASM flux on the CT1 flux , which is given by an exponent of b=1.0$\pm$0.2. The chisquare exhibits a large value since the correlation coefficient has a value of smaller than one (0.83 +- 0.06).



(a) Original ASCA measurement

(b) Subtracted background

Figure B.14: These plots show multiwavelenth measurements of Mkn 421 for the period between April 23, 1998 and May 1, 1998. a) The upper left hand plot shows TeV measurements by Wipple, HEGRA and CAT. The middle left hand plot shows x-ray measurements by ASCA. b) In the upper right hand plot the x-ray measurement is displayed with a background fit and the lower right hand plot shows the x-ray measurement with subtracted background and with superimposed TeV data points for illustration.

Remarks / bin size	Time lag	Fit range
1 day	-0.5 +- 0.39 days	+- 2 days
0.5 day	-0.45 +- 0.15 days	+- 2 days
0.25 day	-0.21 +- 0.15 days	+- 2 days
0.2 day	-0.35 +- 0.15 days	+- 2 days
0.1 day	-0.25 +- 0.1 days	+- 2 days
1 hour	-0.41 +- 0.04 days	+- 2 days
0.5 hour	-0.5 +- 0.04 days	+- 2 days
no bin / interpolation	-0.38 +- 0.07 days	+- 2 days
jacuzzi / 1 hour bin	-0.36 +- 0.1 days	+- 2 days
jacuzzi / interpolation	-0.45 +- 0.078 days	+- 2 days

Table B.4: The table below displays the fit results for different bin sizes. All measurements show a systematic time lag of x-rays of approximately 10 hours. As a crosscheck for the analysis program also the lightcurve of 'jacuzzi', a completely different lightcurve program, written by Daniel Kranich in Munich, was taken to calculate the correlation. It shows the same result. To crosscheck that the effect is not an artifact of binning the correlation was also calculated without binning but by interpolating between neighbouring CT1 measurements.

The analysis is based on the idea that the x-ray flares consist of ***two components***: a ***slow flare*** background component and a ***fast flare*** component which is faster than 0.5 days. The objective in this case is to determine whether TeV flares ***are correlated to the total x-ray*** flux or ***only to the fast flare component***. Therefore, the empirical correlation function is calculated for both the total x-ray flux and for the fast x-ray flares for which the ***slow component has been subtracted*** (see Fig. B.14). The slow component was estimated by fitting the slow flare (of approximately seven day duration) with a Gaussian distribution function.

Figures B.15 Fig. B.16 display the results. The highest correlation for both datasets, with and without background, is achieved for a ***small time lag of a few hours*** (gamma-rays arrive earlier). The effect has only one sigma and is not ***significant.***

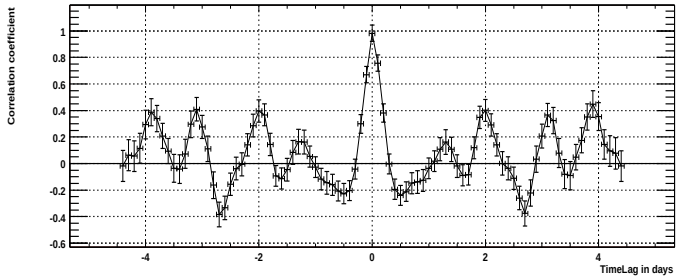
The correlation ***increases*** from 0.63 +- 0.18 (3.5 σ) to 0.70 +- 0.16 (4.3 σ) when the slow x-ray component (i.e. the background) is subtracted. The increase is ***not significant***.

Time lag between the γ-light flux and the x-ray light flux

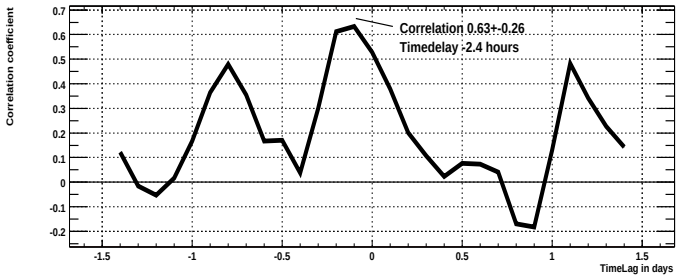
For systematic studies of the time lag the data from CT1 in 2001 were taken. If one looks at Fig. B.17 which used a ***binning of one 1 hour*** and displays time lags Δt of +-6 days one gains the impression that the curve is slightly shifted to the left. This would imply that ***γ-rays arrive slightly earlier than x-rays***.

To clarify the question about the time lag, a Gauss-distribution was fitted to the region around the origin of the correlation plot. First the time lag was determined for different bin widths. Figure B.18 has a bin width of one hour. The fit gives a ***time lag for the x-rays of 9.8 +- 0.9 hours***. Tab. B.4 displays the fit results for ***different bin widths***.

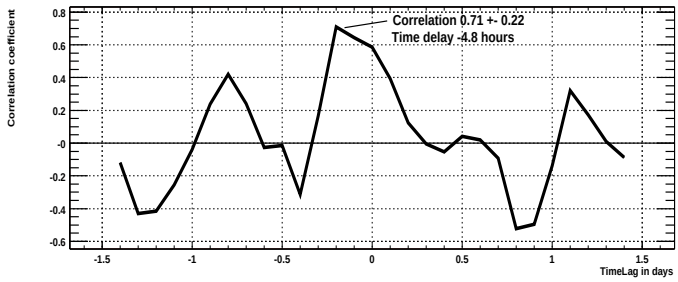
To cross check that this effect is ***not an artifact of binning***, the correlation was also calculated ***without binning***. Since the measurements from ASM/RXTE and CT1 do not precisely coincide, for each 90 s measurement of ASM/RXTE, the value from CT1 was ***interpolated*** between two measurements, if the points were not further apart than 2.5 hours. If there is no CT1 measurement which can be properly interpolated, the measurement is discarded. To ensure that the single correlation points are not correlated, the step size was also chosen to be 2*2.4=4.8 hours (24/5 h). Fig. B.19 shows the fit to the correlation when calculated with interpolated CT1 fluxes. The fit gives a ***time lag of 9.1 +- 1.9 hours***. The range for the fit has been chosen to be +-2 days.



(a) ASCA Autocorrelation



(b) Correlation with slow component



(c) Correlation subtracted slow component

Figure B.15: Plot a) shows an ASCA x-ray autocorrelation plot that illustrates the periodicities in the dataset. Plot b) shows the empirical correlation function before background subtraction which is maximally 0.63 +- 0.18. Plot c) shows the correlation function after background subtraction. It increases to 0.71 +- 0.16 which is insignificant. The two lower plots are very similar if smoothed. The plots b) and c) show a slight tendency for a time shift of a few hours.

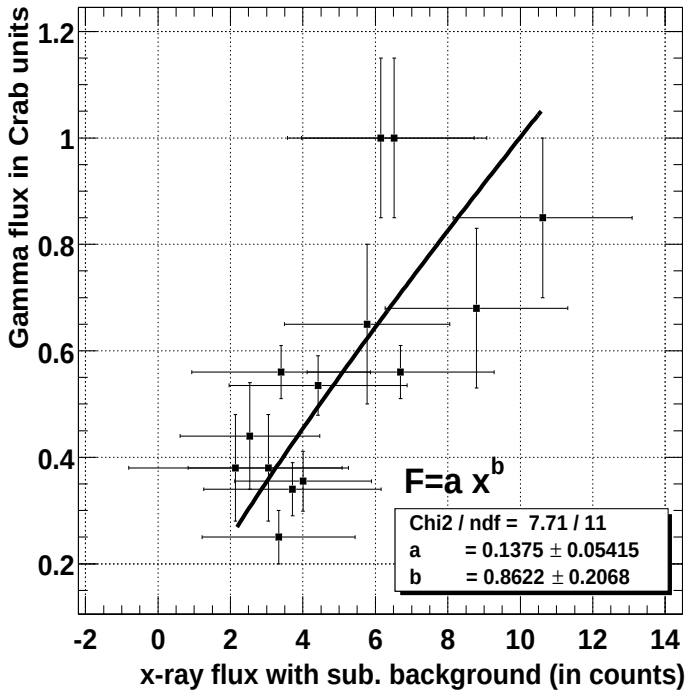


Figure B.16: This plot shows the TeV-gamma Fluxes and the x-ray fluxes plotted against each other. When subtracting the slow flare component from the total x-ray flux it was ensured that the origins of TeV flux and x-ray flux coincide. The fit of $F = ax^b$ reveals an almost perfect linear correlation between the two datasets (b=0.86+-0.2).

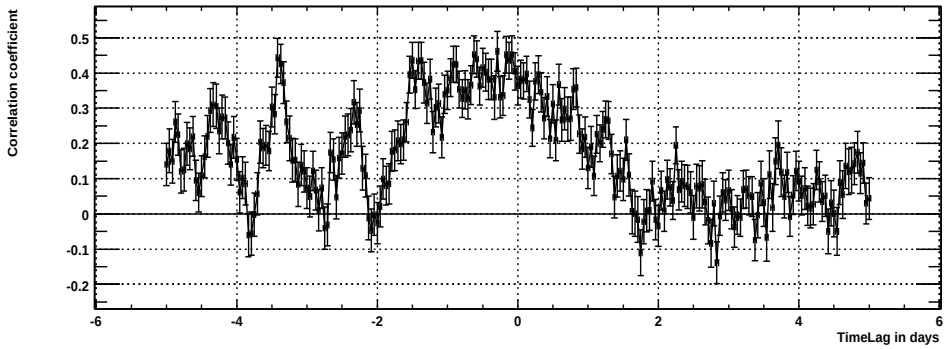


Figure B.17: In this plot with a binning of one 1 hour and time lags Δt from -6 days to +6 days one gets the impression that the curve is slightly shifted to the left. Which would mean that γ***-rays arrive slightly earlier than x-rays.***

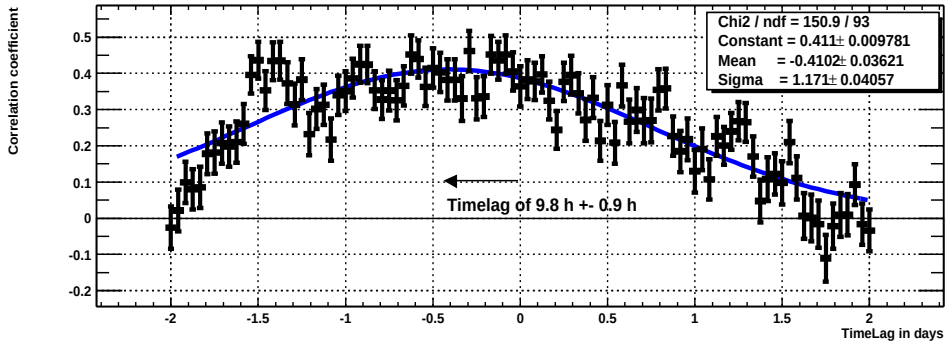


Figure B.18: In this figure only a small range (between +- 2 days) is plotted. The bin size is only 1 hour. The fit shows a time lag of 9.8 hours +- 0.9 hours.

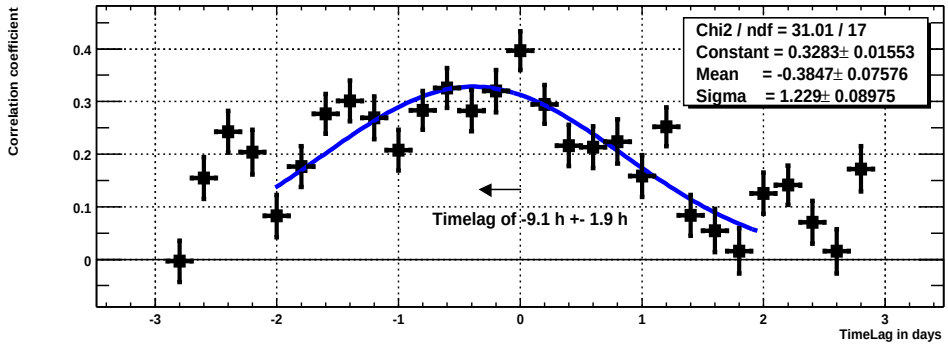


Figure B.19: In this figure the correlation was calculated without binning, but by interpolating between neighbouring CT1 measurements which are not further apart than 2.5 hours. To ensure that the data points are not correlated the step size of the time shift was chosen to be 4.8 hours (24/5 h) as well. The fit gives a time lag of 9.1 hours +- 1.9 hours.

All the measurements with different bin sizes are consistent within the errors and show a time lag of the x-rays of approximately $10 \pm 1_{stat} \pm 3_{sys}$ hours. Since the shape has large fluctuations, a change in the fit range alters the fitted time lag considerably. This means that this result also depends on the fit range. Therefore the ***systematic uncertainty*** is estimated to be approximately ***3 hours***. Taking this into account, the effect is not much larger than approximately three to four sigma for small bin sizes.

The data, especially that from the ASM/RXTE, is too noisy to give an answer of greater precision. Yet, it can be stated that in all of the plots, without exception, a ***systematic shift*** to the left (γ-photons arrive a few hours earlier) has been observed. The result has been cross checked with the light curve calculated by 'jacuzzi' (written by Daniel Kranich in Munich), another light curve calculation program of a completely separate analysis software.

Still, it ***cannot be clarified*** whether the time lag in the discrete correlation function is due to a ***real physical time lag***. This subject will be discussed in detail in chapter 4.

B.1.5 Conclusion about the γ-ray/x-ray correlation studies

CT1-γ-flux and ASM-x-ray-flux:

It can be concluded that there is a ***significant correlation*** of 0.74 +- 0.05 (in 1 day bins) between the x-ray flux and the γ-ray flux. The correlation between these fluxes can clearly be seen even by eye in Fig. B.13. A fit of a function $F = ax^b$ shows that the ***correlation is almost linear*** (b=1.0+-0.2). The empirical correlation coefficient increases monotonically with bin size probably due to the improved statistics of the bins that are involved. It is possible that this increase with bins size has a physical reason as well***.***

The hint about a possible ***time lag*** between the x-rays and the gammas of a few hours (~$10 \pm 1_{stat} \pm 3_{syst}$ hours) is of considerable interest. This time lag has not only been observed with the CT1 lightcurve calculated by this program package in C++/ROOT, but also using the lightcurve calculated by the ***'jacuzzi'*** program written by Daniel Kranich in Munich. This proves redundancy of the lightcurve calculation. The apparent time lag of the x-rays can be also explained with an asymmetric shape of the (longer) flares. Different parameters for the rise (acceleration) and fall time (cooling) could result in such an effect.

ASCA-x-ray measurement:

The ASCA x-rays measurement from a flare of Mkn 421 in April/May 1998 is much more detailed than the ASM measurements. The x-ray flares can be ***divided into a slow flare component*** (of ~seven days duration) ***and a very fast flare component*** (shorter than 0.5 days duration). Also for the measurement from ASCA a ***clear correlation*** between x-ray flux and TeV γ-flux can be observed. It was studied whether the γ-flux is mainly correlated to the fast x-ray component or to the total x-ray flux. For both cases a rather high correlation was found. The maximum correlation for the TeV flux and the total x-ray flux is 0.63 +- 0.22. If the s***low component is subtracted*** the correlation ***increases*** to 0.71 +- 0.16 which is ***insignificant.***

The correlation between the TeV γ-flux and x-ray-flux is almost linear just as above (b=0.85 +-0.2) and, as before, a weak hint of a time lag between x-rays and γ's of approximately 4 hours was be observed.

The observed timelag:

The observed timelag of the x-rays has only an estimated significance of ***three sigma*** since the said systematical error is large. However, a potential timelag could originate from either ***opacity effects*** (x-rays escape later and γ escape immediately from the surface) or ***cooling effects*** of the VHE electrons in the jet (in the beginning many γ-rays are produced but later when the electrons cooled down they only emit x-ray-synchrotron radiation).

The timelag could even be a first sign of quantum gravity effects where theoreticians expect changes in the speed of light for photons with energies close to the Planck mass. The time lag of x-rays would point to an ***increase*** of the speed of light for increasing photon energies. Even that the energy of TeV photons is far away from the Planck scale, the long distance from Mkn 421 to earth could help to amplify even tiny effects.

Still, it ***cannot be clarified*** whether the time lag in the discrete correlation function is due to a ***real physical time lag***. The effect could also be due to a very ***asymmetric shape*** of the x-ray flares. This subject will be discussed in chapter 4.

B.1.6 The hardness ratio as measured by CT1 in 2001

The hardness ratio describes the ***steepness*** of a differential spectrum for a given source (e.g. Mkn 421) in small ***time bins*** (which correspond to single runs in our case). It is calculated together with the lightcurve as was described in detail earlier. Its ***advantage*** over a spectrum is that it ***needs less statistics*** and can be calculated for single runs. The hardness ratio gives us the ***possibility*** to see whether the spectral index of the power law spectrum ***changes*** during different flare states. Figures B.20, B.21 and B.22 show the hardness ratio for different energy intervals as a function of the total flux.

The hardness ratio is the ratio of the integrated flux of a higher energy interval and that of a lower energy interval. For an assumed power law spectra, $\frac{dF(t')}{dE} = F_0\,(t)\,E^{-\alpha}$, there is a clear relation between the hardness ratio and the spectral index. The hardness ratio for a time bin i and the energy intervals $[E_{upper,min}, E_{upper,max}]$ and $[E_{lower,min}, E_{lower,max}]$ is given by:

$$H\,(t_i) = \frac{\int_{t_{min,i}}^{t_{max,i}} dt' \int_{E_{upper,min}}^{E_{upper,max}} dE\, \frac{dF(t')}{dE}}{\int_{t_{min,i}}^{t_{max,i}} dt' \int_{E_{lower,min}}^{E_{lower,max}} dE\, \frac{dF(t')}{dE}} \tag{B.4}$$

$$= \frac{\frac{F_{0,i}\left[E^{-\alpha_i+1}\right]_{E_{upper,min}}^{E_{upper,max}}}{-\alpha_i+1}}{\frac{F_{0,i}\left[E^{-\alpha_i+1}\right]_{E_{lower,min}}^{E_{lower,max}}}{-\alpha_i+1}}$$

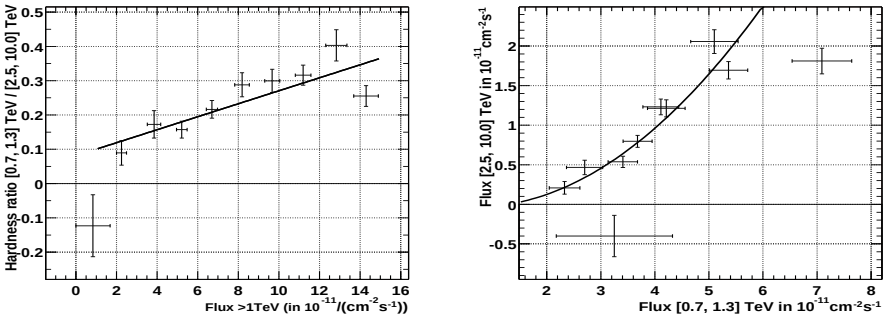
$$H\,(t_i) = \frac{E_{upper,max}^{-\alpha_i+1} - E_{upper,min}^{-\alpha_i+1}}{E_{lower,max}^{-\alpha_i+1} - E_{lower,min}^{-\alpha_i+1}} \tag{B.5}$$

For a given hardness ratio the differential spectral index α_i can be obtained numerically.

To detect a change in the hardness ratio as a function of the flux in the data of the very fast flaring source Mkn 421 the following procedure was applied.

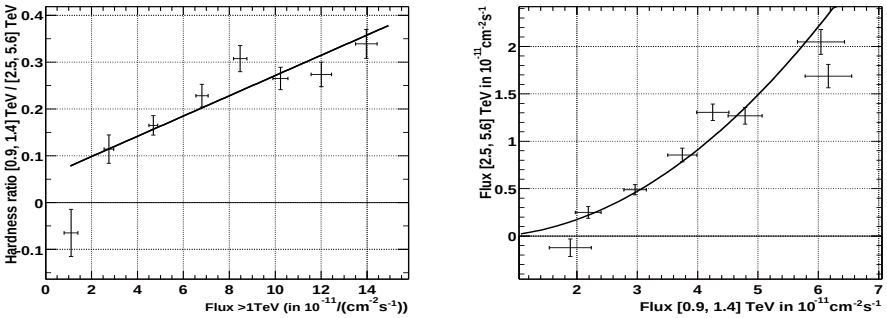
First the fluxes in the lower and upper energy interval were calculated for each run. Unfortunately the statistics of a single 20 minute run does not contain enough statistical information to calculate the hardness within that run. To collect ***more statistics***, the measurements of the fluxes in the low energy interval and the high energy interval were ***re-binned*** into total flux bins (consisting of approximately 10 bins), as described in Equ. E.35 (Appendix A). Since some of the measurements had huge errors, a quality cut of $F_i > \sigma_i$ was introduced. This quality cut did not change the shape of the fit result but it did reduce the fluctuations in the re-bins significantly.

Tab. B.5 ***shows the quantitative fit results*** for different energy intervals (LI=Lower energy interval, UI=upper energy interval). For each plot a constant ($y = a$, denoted as ***'CO'***) and a straight line ($y = a + mx$ denoted as ***'SL'***) were fitted and the quantity χ^2/NDF and its significance of rejection (of the constant fit only) were computed. For a straight line, the parameters (in units of hardness ratio) are 'a' (a constant) and 'm' (the slope). For a constant fit there is just the single parameter 'a'. The last two columns contain the spectral index for an assumed power law spectrum for a flux of $F = 4.0 \cdot 10^{-11} cm^{-2} s^{-1}$ and for $F = 12.0 \cdot 10^{-11} cm^{-2} s^{-1}$.



(a) (b)

Figure B.20: These plots display the hardness ratio for the intervals [0.7, 1.3] TeV and [2.5, 10.0] TeV and the correlation between the flux in the lower energy interval and the higher energy interval. a) The ***hardness ratio clearly changes with flux***. A constant fit was rejected with 4.2 sigma. In figure b) the flux in the lower energy interval [0.7, 1.3] TeV is plotted against the flux in the upper energy interval [2.5, 10.0] TeV as a cross check. This is a plot to cross check the distribution. The χ^2/NDF values are listed in Tab. B.5.



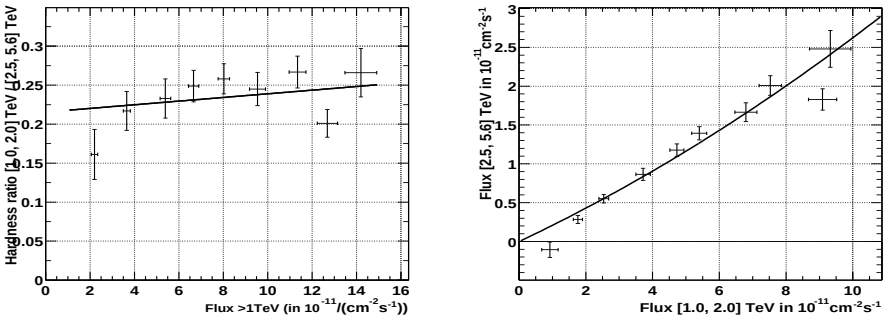
(a) (b)

Figure B.21: a) This plot displays the hardness ratio for the intervals [0.9, 1.3] TeV and [2.0, 5.6] TeV. It shows that the ***hardness ratio again increases with flux***. Although the effect is smaller than that of the previous plot, a constant fit was rejected with 3.3 sigma. In figure b) the flux in the lower energy interval ([0.9, 1.4] TeV) is again plotted against that in the upper energy interval ([2.0, 5.6] TeV), as a cross check. The χ^2/NDF values are listed in Tab. B.5.

LI/TeV	UI/TeV	Fit	$a/10^{-2}$	$m/10^{-2}$	χ^2/NDF	Sig
[0.7, 1.3]	[2.5, 10]	CO	24±4		57.3/9	4.2
[0.7, 1.3]	[2.5, 10]	SL	8.1 ± 2.7	1.9 ± 0.3	17.5/8	
[0.9, 1.4]	[2.5, 5.6]	CO	17±4		83.5/7	5.6
[0.9, 1.4]	[2.5, 5.6]	SL	5.5 ± 2.3	2.1 ± 0.3	12.6/6	
[1.0, 2.0]	[2.5, 5.6]	CO	24±3		15/9	1.2
[1.0, 2.0]	[2.5, 5.6]	SL	22 ± 2.0	0.2 ± 0.2	13.9/8	
[1.6, 2.2]	[2.5, 5.6]	CO	80±7		9.47/12	0.4
[1.6, 2.2]	[2.5, 5.6]	SL	80 ± 7.0	0 ± 1.0	9.47/11	

LI/TeV	UI/TeV	Fit	$\alpha(Flux = 4.0)$	$\alpha(Flux = 12.0)$
[0.7, 1.3]	[2.5, 10]	CO	2.5 ± 0.1	
[0.7, 1.3]	[2.5, 10]	SL	2.7 ± 0.1	2.3 ± 0.1
[0.9, 1.4]	[2.5, 5.6]	CO	2.75 ± 0.15	
[0.9, 1.4]	[2.5, 5.6]	SL	3.1 ± 0.15	2.4 ± 0.15
[1.0, 2.0]	[2.5, 5.6]	CO	2.65 ± 0.2	
[1.0, 2.0]	[2.5, 5.6]	SL	2.68 ± 0.3	2.62 ± 0.3
[1.6, 2.2]	[2.5, 5.6]	CO	2.8 ± 0.2	
[1.6, 2.2]	[2.5, 5.6]	SL	2.8 ± 0.2	2.8 ± 0.2

Table B.5: The table shows ***the quantitative fit results*** for different energy intervals (LI=Lower energy interval, UI=upper energy interval). For each plot a constant ($y = a$, denoted as ***'CO'***) and a straight line ($y = a + mx$ denoted as ***'SL'***) were fitted and the quantity χ^2/NDF and its significance of rejection (of the constant fit only) were computed. For a straight line, the parameters (in units of hardness ratio) are 'a' (a constant) and 'm' (the slope). For a constant fit there is just the single parameter 'a'. The last two columns contain the ***spectral index*** for an assumed ***power law*** spectrum for a flux of $F = 4.0 \cdot 10^{-11} cm^{-2} s^{-1}$ and for $F = 12.0 \cdot 10^{-11} cm^{-2} s^{-1}$.
It can be seen that the hardness ratio increases with increasing flux but only for the lower energy interval just above the threshold. There, the constant fit was ***rejected with 4.2 sigma***.



(a) (b)

Figure B.22: Plot a) shows the hardness ratio for the intervals [1.0, 2.0] TeV and [2.5, 5.6] TeV. The change in hardness ratio becomes much smaller for higher energies. In the region above 2 TeV the ***hardness ratio is almost constant*** but still displays a slight increase with increasing flux. Figure b) displays the cross check for the intervals [1.0, 2.0] TeV and [2.5, 5.6] TeV. Here the correlation plot shows almost a straight line. The hardness ratio is virtually constant. The χ^2/NDF values are listed in Tab. B.5.

It can be seen that the hardness ratio increases with increasing flux but only for the lower energy interval just above the threshold. There, the constant fit was ***rejected with 4.2 sigma.*** Lower energy intervals above 1.6 TeV show a constant hardness ratio. This will be discussed in detail in the conclusions below.

Conclusion concerning the hardness ratio calculation

In the figures above it could already been seen that ***the hardness ratio increases with increasing flux***. However, the change ***increases*** with ***decreasing*** energy. The effect is largest just above the threshold. Above 2.0 TeV the hardness ratio stays virtually constant. To determine whether this is a significant effect, potential systematic errors must be discussed carefully.

Energy intervals close to the threshold introduce large systematic errors in the calculation of the hardness because the slope of the effective detection area is very steep in those regions. However, these ***systematic errors in the calculation of the hardness are static***, they do not change with time or run. To ***exclude any bias*** from varying thresholds due to different zenith angles ***only events with zenith angles between 10° and 20°*** were considered. The threshold remains virtually unchanged in this region. Even if there are systematic errors in the flux calculation (and thus in the hardness ratio calculation), due to incorrect collection areas too close to the threshold, the effect of a ***change*** in the hardness ratio will persist.

The ***spectral indices***, obtained from the hardness ratios, ***are consistent with the spectrum*** which is presented in the following section. In general, they clearly show that the spectral index (at energies of approximately 1 TeV) is ***increasing from about*** $\alpha = 2.4$ ***up to about*** $\alpha = 3.0$ from low fluxes ($F = 4.0 \cdot 10^{-11} cm^{-2} s^{-1}$) to high fluxes ($F = 12.0 \cdot 10^{-11} cm^{-2} s^{-1}$).

How could a different hardness ratio behavior for different energy intervals be explained ? An explanation could be that the peak emission of inverse Compton scattered γ-photons is not far away from the measurement (lets say at about 10 to 100 GeV). Below 1 TeV the spectrum does not have a power-law shape anymore. It is curved there and only above 1 TeV it gets a power law shape. During flare the peak moves slightly to higher

energies. The hardness ratio only changes if parts of the lower energy interval is below approximately 1 TeV. This result will be compared with observations from the spectrum.

B.2 Spectrum of Mkn 421

The last aspect of the Mkn 421 data sample, that is going to be analyzed is the shape of the spectrum. The ***energy bin size*** (in logarithmic scale) has been chosen to be of the double size (two sigma) of the average energy resolution (~23 %).

In the second stage the spectra of different flux-states are compared to each other. As predicted by the studies of the hardness ratio a change in the slope of the spectrum at lower energies is expected.

B.2.1 The spectrum averaged over all flares

The average spectrum was calculated out of the complete dataset of the flares of Mkn 421 in 2001 (which accumulate to 259 hours of observation time).

Power law with exponential cutoff hypothesis

Figure B.23 shows the average spectrum of Mkn 421 for the period between 59125 MJD and 59250 MJD. Two differential flux $\frac{F(E)}{dE}$ functions have been fitted to the spectrum.

1. The first hypothesis is a simple ***power-law spectrum*** with spectral index α:

$$\frac{F\,(E)}{dE} = F_0 E^{-\alpha} \tag{B.6}$$

2. The second hypothesis is a ***power-law with an exponential cutoff*** E_c:

$$\frac{F\,(E)}{dE} = F_0 E^{-\alpha}\, e^{-\frac{E}{E_c}} \tag{B.7}$$

The first hypothesis of a power-law spectrum gave a spectral index of α =2.96 +- 0.04 but has been ***rejected*** with a chi-square of χ^2/NDF=74/9. This has a probability of p=10^{-21} or 6.6 sigma (Gaussian normal distribution).

The spectrum has been fitted with a power law having an exponential cutoff (upper plot). The ***spectral index is*** α ***=1.88 +- 0.15*** and the ***cutoff is located at*** E_c ***=3.1 +- 0.5 TeV***. The points with ***different colors represent 6 different zenith angle bins*** ranging up to 50°. The ***black points*** are data from ***all zenith angles*** combined. In the lower left hand plot the expected excess event distribution for a power law with cutoff has been fitted to the experimental excess event distribution. The ***fit to the excess event distribution*** is expected to yield the preciser results because the unfolding procedure introduces additional errors. The spectral index found here is α ***=2.1 +- 0.07 with a cutoff of*** E_c ***=3.1 +- 0.26 TeV*** and is therefore perfectly consistent with the upper plot. The plots for the background estimation for each (unfolded) energy bin can be found in ***Appendix B***. The background for the measurement at 12 TeV was difficult to estimate and could have a higher systematic uncertainty.

The systematic uncertainties, as already discussed in the last chapter, have their origin mainly in imperfect calibration of the absolute energy scale and in imperfect MC studies. It is estimated that the systematic uncertainty of the spectral index is $\sigma(\alpha_{sys}) = 0.1$, that of the energy cutoff is $\sigma(E_{cut}) = 1\,TeV$ and that of the flux is $\sigma(F_0) = 10\%\,F_0$.

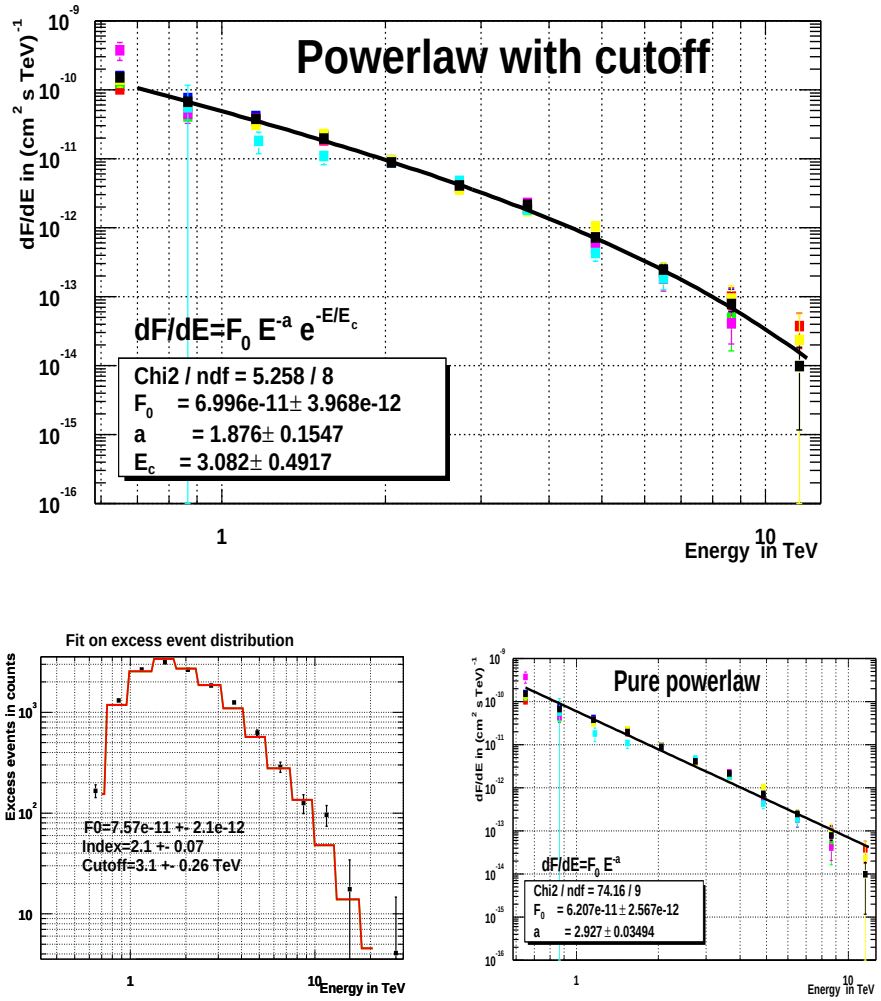


Figure B.23: Plot a) shows the average spectrum. It was fitted with a power law having a exponential cutoff. The ***spectral index is 1.88 +- 0.15*** and the ***cutoff is located at 3.1 +- 0.5 TeV.*** The points with different colors represent 6 different zenith angle bins ranging up to 50°. The black points are data from all zenith angles combined. In plot b) the expected excess event distribution for a power law with cutoff was fitted to the ***experimental excess event distribution***. This avoids the unfolding procedure and yields smaller errors for the fit parameters. The ***spectral index found here is 2.1 +- 0.07*** with a ***cutoff of 3.1 +- 0.26 TeV.*** Plot c) shows the fit of a pure power law fit which is rejected with a χ^2/NDF=74/9.

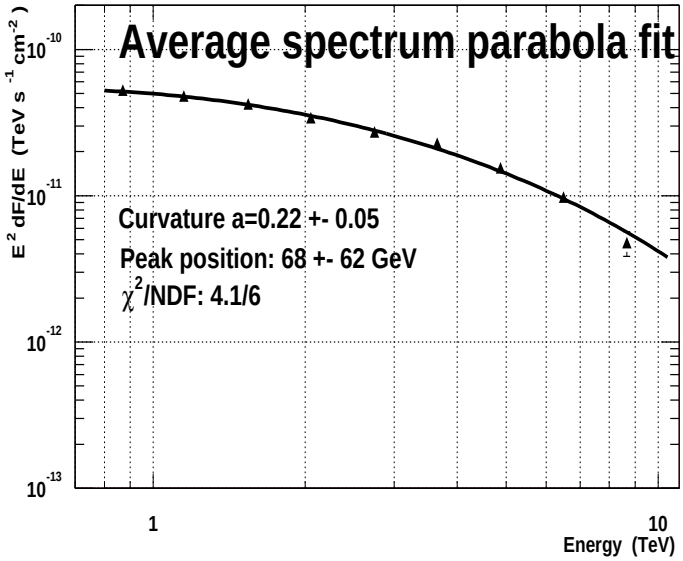


Figure B.24: The curvature in the spectrum shown here might not only originate from a cutoff, but from an intrinsic curvature of the spectrum, since what is observed, is supposed to be a portion of the inverse Compton peak. A function that is proportional to the luminosity and contains a parabolic term was fitted here: $L \propto E\frac{dF}{dE} = F_0 E^{-\alpha-\beta \ln E}$. The chisquare is χ^2/NDF=4.1/6, which is perfectly acceptable. The fit gives a value for the curvature term of a=0.22, which will be used later when the spectra of different flux states are analyzed. The spectrum was plotted in units of $E^2 dF/dE$ for better illustration.

Power law with a parabolic term

The observed curvature in the spectrum might ***not only*** originate from ***a potential cutoff*** (for example from the absorption of the cosmic infrared background) but also from an ***intrinsic curvature*** of the spectrum. What we are supposed to observe is a portion of the ***inverse Compton peak,*** which naturally posses a curvature rather than a pure power law shape. This idea will be studied further when the spectra of different flux levels are examined.

To first order the peak shape can be approximated by a ***quadratic term***. The resulting function is a parabola (in log-log scale). Usually SSC model predictions are presented in terms of ***luminosities***. For this reason the ansatz for the fit function is ***proportional to the luminosity*** $L \propto E\frac{dF}{dE}$. Naturally, the quadratic term is introduced in log-log scale and the parabola ansatz becomes:

$$
\begin{aligned}
E\frac{dF}{dE} &= e^{-a(\ln E - \ln E_{peak})^2 + c} && \text{(B.8)} \\
&= e^{c} \cdot e^{-a \ln^2 E - 2a \ln E \ln E_{peak} + -a \ln^2 E_{peak}} \\
&= e^{c} \cdot e^{-a \ln E_{peak}} \cdot E^{+2a \ln E_{peak} - a \ln E} \\
&= F_0 \cdot E^{-\alpha-\beta \ln E} && \text{(B.9)}
\end{aligned}
$$

The last line shows that the ansatz is equal to a usual power law with a quadratic term β if

$$
\begin{aligned}
F_0 &= e^{c} \cdot e^{-a \ln E_{peak}} && \text{(B.10)} \\
\alpha &= 2a \ln E_{peak} \\
\beta &= a
\end{aligned}
$$

Instrument	Cangaroo	Whipple	CAT
Years	1992-1996	1995-1998	1996-2000
Range [TeV]	7-50	0.5-5	0.3-20
F_0	$2.8^{+1.9}_{-1.2}$	$3.20 \pm 0.17 \pm 0.6$	$2.2 \pm 0.05 \pm 0.6$
α	2.53 ± 0.18	$2.49 \pm 0.06 \pm 0.04$	$2.80 \pm 0.03 \pm 0.08$
Reference	[Tan98]	[Hil98]	[Mas01]

Instrument	CT-System	CT1
Years	1997-1999	2001
Range [TeV]	0.5-20	0.7-20
F_0	$2.79 \pm 0.02 \pm 0.5$	$3.0 \pm 0.33 \pm 0.5$
α	$2.59 \pm 0.03 \pm 0.05$	$2.5 \pm 0.1 \pm 0.1$
Reference	[Aha00]	[TW]

Table B.6: The table lists the actual flux measurements of the Crab nebula with different instruments.

Fig. B.24 shows that this ansatz fits quite well to the shape of the spectrum. The ***chi-square is*** χ^2***/NDF=4.1/6***. The fit gives us a value for the ***quadratic term of a=0.22,*** which will be used later when the spectrum for different flare states is examined. The fit also provides us a value for the peak of the parabola. Since the shape of the inverse Compton peak is not known and since SSC models do not predict a parabolic inverse Compton peak, this value only provides a very rough hint of the position of the real luminosity peak.

B.2.2 The Crab nebula spectrum for consistency cross check

The Crab nebula is a pulsar that permanently feeds a shock front which produces γ-rays. The spectra of shock wave acceleration have a power law shape. It has become a standard candle or calibration light source since it emits a time constant flux with a ***power law*** spectrum, given by:

$$\frac{dF}{dE} = F_0 E^{-\alpha} \tag{B.11}$$

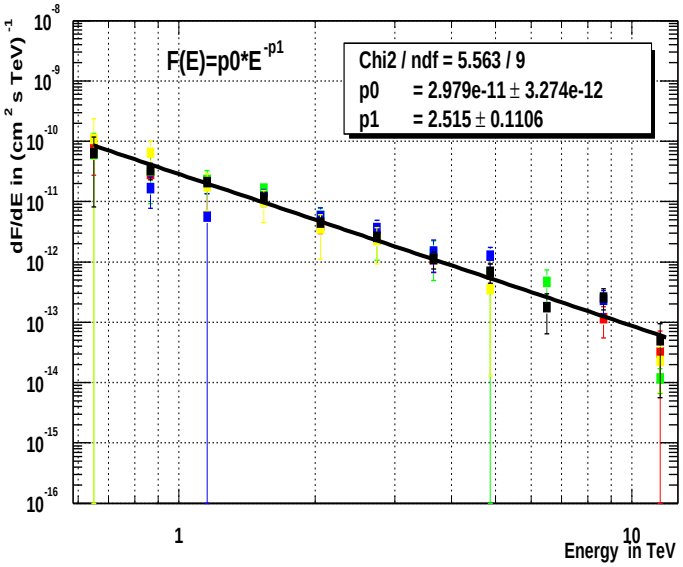
where α is the spectral index and F_0 is the flux constant. The spectrum of the Crab Nebula was measured by many experiments. It can be used to check if the one's own analysis is consistent with common accepted results.

Fig. B.25 shows the spectrum as recorded by CT1 during the same period as Mkn 421 in ***2001***. The spectrum is flatter (or harder) than that of Mkn 421. Since the Crab nebula is in our galaxy no absorption from the cosmic infrared background is expected. The ***spectral index, as measured in this work, is*** $\alpha = 2.5 \pm 0.1 \pm 0.1$ ***and the flux constant*** $F_0 = 3.0 \pm 0.33 \pm 0.5\, TeV^{-1} cm^{-2} s^{-1}$. The energy range is, as can be seen in the figure, 650 GeV to 15 TeV. The second figure below shows the zenith angle dependence of the flux which is negligible (as is should be). Tab. B.6 shows the results from several Cherenkov telescopes.

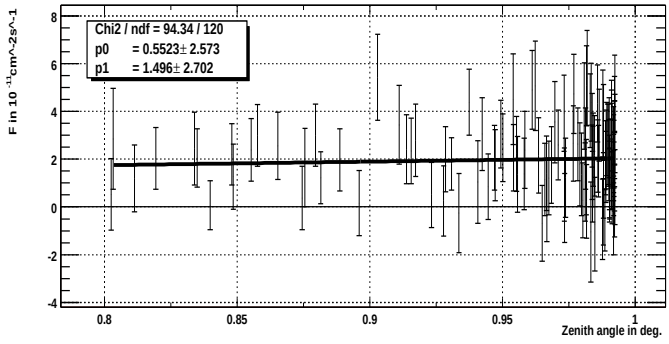
B.2.3 Analysis of the Mkn 421 spectrum during different flare states

Remembering the results from the analysis of the hardness ratio we expect a change of curvature in the spectrum in the region at approximately 1 TeV and eventually a slight change of the spectral index of the spectrum in dependence to different flare states.

The data has been separated in essentially ***five intervals of different fluxes*** ($[0.5, 1.0]$, $[6.0, 10]$ and $[10, 15]$ in $10^{-11} TeV^{-1} cm^{-2} s^{-1}$). For all of these flare states the spectrum has been calculated.



(a) Crab Spectrum



(b) Flux vs zenith angle with straight line fit

Figure B.25: The Crab nebula has become a calibration light source for Cherenkov telescopes since it emits a constant flux of γ-rays with a power law spectrum. Plot a) shows the spectrum when fitted by a power law. As usual, different colors represent different zenith angles bins and the black points are the sum of all zenith angles combined. The fit gives a ***spectral index of*** $\alpha = 2.5 \pm 0.1$ and a ***flux constant of*** $F_0 = 3.0 \pm 0.33$. b) A straight line fit ($y = p0 + p1 * x$) shows that the zenith angle dependence of the flux above 1 TeV is negligible.

$Interval$	T_{obs}	F_0	α	χ^2/NDF
in TeV	in h	$TeV^{-1}cm^{-2}s^{-1}$		Fixed cutoff, free slope
$[0.5, 1.0]$	21.7	$2.05 \pm 0.26\,10^{-11}$	2.50 ± 0.2	$1.0/4$
$[1.0, 2.5]$	82.9	$3.17 \pm 0.17\,10^{-11}$	2.22 ± 0.08	$2.3/5$
$[2.5, 4.5]$	79.8	$7.07 \pm 0.28\,10^{-11}$	2.04 ± 0.04	$3.9/5$
$[7.0, 10.0]$	23.9	$13.4 \pm 0.53\,10^{-11}$	1.88 ± 0.04	$2.8/5$
$[10.0, 15.0]$	11.8	$19.0 \pm 8.13\,10^{-11}$	1.83 ± 0.05	$5.8/5$

Table B.7: A power law with fixed cutoff of $E_c = 3.2\,TeV$ (result from overall spectrum) was fitted. For lower fluxes the spectrum becomes ***steeper.***

$Interval$	T_{obs}	F_0	χ^2/NDF
in TeV	in h	$TeV^{-1}cm^{-2}s^{-1}$	Fixed cutoff, fix slope
$[0.5, 1.0]$	21.7	$2.05 \pm 0.26\,10^{-11}$	$17.8/5$
$[1.0, 2.5]$	82.9	$3.17 \pm 0.17\,10^{-11}$	$30.7/6$
$[2.5, 4.5]$	79.8	$7.07 \pm 0.28\,10^{-11}$	$25.6/6$
$[7.0, 10.0]$	23.9	$13.4 \pm 0.53\,10^{-11}$	$4.0/6$
$[10.0, 15.0]$	11.8	$19.0 \pm 8.13\,10^{-11}$	$5.8/6$

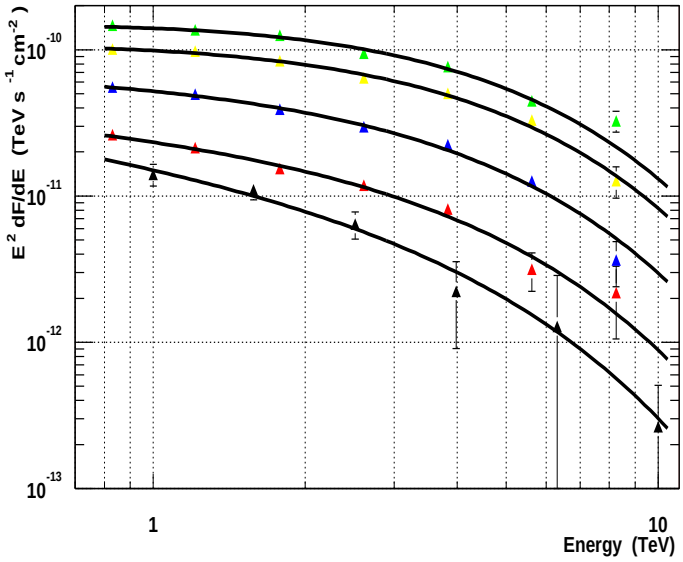
Table B.8: A power law with fixed cutoff of $E_c = 3.2\,TeV$ and fixed slope (slope of highest flux curve) was fitted. ***The fixed slope hypothesis was rejected by χ^2/NDF=30.7/6*** (maximum value) or ***3.7 sigma.***

Power law with exponential cutoff

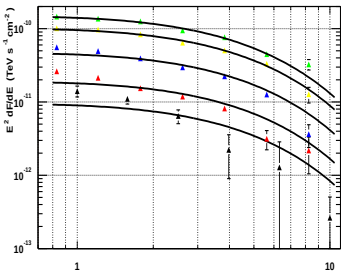
The first hypothesis tested is a power law with cutoff. Figure B.26 shows the $E^2 \frac{dF}{dE}$ spectra plots of these five flux intervals.
a) A power law with fixed cutoff of $E_c = 3.2\,TeV$ (result from overall spectrum) has been fitted. It was necessary to fix one parameter to reduce the errors on the fit parameters to obtain meaningful results. The spectral index of the different flare states changes considerably. ***For lower fluxes the spectrum becomes steeper***. The fit results are listed in Tab. B.7.

b) To check the significance of the change in slope during different flares ***the slope was fixed*** to $\alpha = 1.8$, that with the highest flux (green points) and the χ^2 was obtained from each fit. ***The fixed slope hypothesis was rejected by χ^2/NDF=30.7/6*** (maximum value) or ***3.7 sigma*** (for a Gaussian normal distribution). Tab. B.8 shows that none of the other four curves fit the fixed shape well.

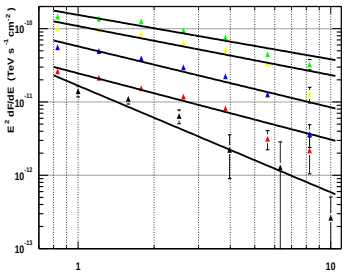
c) In the first fit 'a)' the cutoff has been fixed and the spectral index was free. Now we check if the fit is also compatible with a ***fixed slope of*** $\alpha = -2.0$ (slope of overall, average spectrum) and a ***free cutoff***. The fit proves that the data is ***also perfectly compatible*** with a fixed slope and a free cutoff. The ***cutoff*** moves to ***higher*** energies with ***increasing*** flux. The information available is not sufficient to differentiate between the different models. Tab. B.9 lists the fit results.
d) To check the significance of the ***existence*** of a ***cutoff*** (or a curvature) a ***pure power law*** was also fitted. ***The pure power law was rejected by a chisquare of χ^2/NDF=49/5*** (maximum value).The fit results are listed in Tab. B.10. The values in the table illustrate how the spectrum ***hardens*** with increasing flux.



(a) Powerlaw with fix cutoff



(b) Lowerlaw with fix cutoff and fix slope



(c) Pure powerlaw fit

Figure B.26: These are $E^2 \frac{dF}{dE}$ spectra plots of various flare states. The data was separated into five intervals of different fluxes. a) The spectra for each interval was calculated and a power law with a ***fixed cutoff of*** $E_c = 3.2\,TeV$ was fitted to the data. To obtain meaningful results it was necessary to fix one parameter so that the errors on the fit parameters are reduced. The spectral index of the different flare states changes considerably. In the plot b), to check the significance of the change in slope during different flares, the slope has was to the one with the highest flux (the green points) and the χ^2 was obtained from the fits. The fixed slope was rejected by χ^2***/NDF=30.7/6*** (maximum value), or ***3.8 sigma***. c) To check the significance of the existence of a cutoff a pure power law was fitted to the data. The pure power law was rejected by χ^2***/NDF=49/5*** (maximum value).

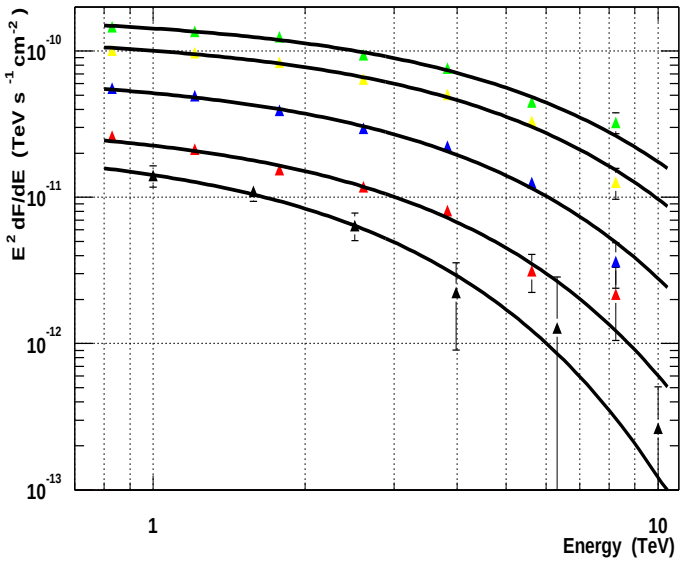


Figure B.27: This figure shows the fit of a power law with a fixed spectral index of $\alpha = -2.0$ and a free cutoff. The χ^2-values also agree to such a solution (See Tab. B.9). The cutoff moves from 1.9 TeV up to 4.3 TeV. The data is not good enough to distinguish between these different possibilities.

$Interval$	T_{obs}	F_0	Cutoff	χ^2/NDF
in TeV	in h	$TeV^{-1}cm^{-2}s^{-1}$	in TeV	Free cutoff, fixed slope
$[0.5, 1.0]$	21.7	$2.4 \pm 0.55\,10^{-11}$	1.9 ± 0.4	$0.8/4$
$[1.0, 2.5]$	82.9	$3.4 \pm 0.27\,10^{-11}$	2.5 ± 0.23	$3.5/5$
$[2.5, 4.5]$	79.8	$7.1 \pm 0.35\,10^{-11}$	3.1 ± 0.18	$4.0/5$
$[7.0, 10.0]$	23.9	$13.0 \pm 0.60\,10^{-11}$	4.0 ± 0.25	$2.8/5$
$[10.0, 15.0]$	11.8	$18.1 \pm 0.84\,10^{-11}$	4.3 ± 0.3	$3.4/5$

Table B.9: Here the ***slope was fixed*** and ***cutoff was kept free***. As expected the cutoff moves to ***higher*** energies for ***increasing*** flux. The chisquare values are acceptable. The slope and the cutoff are highly correlated parameters. The information available does not permit to differentiate between different models.

$Interval$	T_{obs}	F_0	α	χ^2/NDF
in TeV	in h	$TeV^{-1}cm^{-2}s^{-1}$		Pure powerlaw
$[0.5, 1.0]$	21.7	$1.7 \pm 0.19\,10^{-11}$	3.50 ± 0.14	$7.4/4$
$[1.0, 2.5]$	82.9	$2.5 \pm 0.12\,10^{-11}$	2.90 ± 0.06	$11.6/5$
$[2.5, 4.5]$	79.8	$5.8 \pm 0.20\,10^{-11}$	2.83 ± 0.04	$49.0/5$
$[7.0, 10.0]$	23.9	$10.9 \pm 0.37\,10^{-11}$	2.66 ± 0.04	$42.5/5$
$[10.0, 15.0]$	11.8	$15.5 \pm 0.58\,10^{-11}$	2.6 ± 0.04	$23.4/5$

Table B.10: To check the significance of the ***existence*** of a ***cutoff*** (or a curvature) also a ***pure power law*** was also fitted. ***The pure power law was rejected by a chisquare of χ^2/NDF=49/5.***

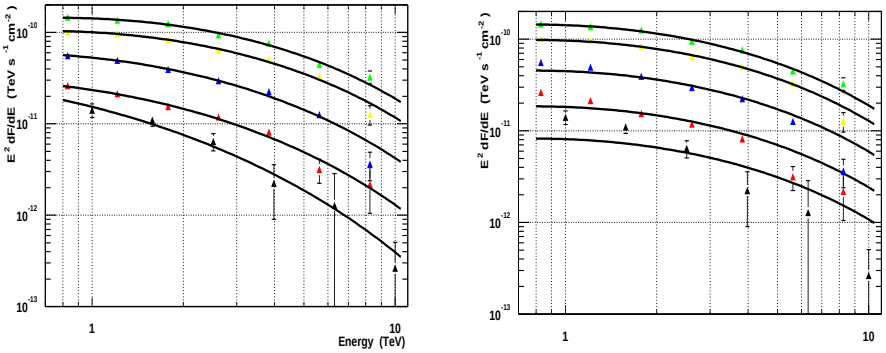


Figure B.28: A ***power law with a quadratic term*** describes the curvature of the spectrum very well. In plot a) Each spectrum has been fitted with a fixed curvature of a=0.22. This value was measured previously from the overall spectrum. From the fit the peak position of the parabola was calculated. ***It moves significantly from 13 GeV to 78 GeV***. Plot b): The peak position was fixed to E_{peak}=78.8 GeV in order to check the significance of the movement. It is evident that the resulting curves do not describe the shape of the lower flux spectra. The fits are ***rejected*** with a chisquare of χ^2***/NDF= 33/6*** (maximum value). This corresponds to a probability of $5\ 10^{-8}$.

Power law with a quadratic (parabolic) term

As already mentioned, the curvature of the spectrum can also be explained also by the fact that one observes a part of the inverse Compton peak. Therefore, it makes sense to perform a fit using a ***power law with a quadratic curvature term***. This means that the inverse Compton peak is approximated by a parabola. SSC models normally give predictions about ***luminosities***. Our ansatz for the fit function is

$$L \propto E\frac{dF}{dE} = I_0 E^{-\alpha-\beta \ln E}$$

where β is the ***curvature***, α is ***spectral index*** and I_0 is the ***peak intensity*** (flux times energy). Using Equ. B.10, the ***peak position of the parabola*** can be calculated. It must be kept in mind that this value is ***not the position of the luminosity peak*** but rather a ***very rough estimate*** since the inverse Compton peak does ***not*** have a parabolic shape! Fig. B.28 shows the spectra for the five different flux intervals to which a power law with quadratic a curvature term was fitted.

What we want to test is whether the change in spectral index (or equivalent the cutoff), can be explained by the ***idea of a moving peak***. The curvature was fixed to a value of ***a=0.22*** as previously measured from the overall spectrum. The hypothesis is that the peak ***only moves*** with increasing fluxes but does ***not change its shape***. The fixing of a is also necessary to limit the number of free fit parameters to two (just as was done for the cutoff fit). This reduces the errors on the fit parameters sufficiently to obtain meaningful results.

The fit results show that the ***peak of the parabola moves from 13 GeV to 78 GeV***. The value for the intensity I_0, which is proportional to the ***peak*** luminosity suggests that the ***luminosity at peak does not change very much***. This simple model seems to show that the increase in flux (in the TeV) is mainly caused by a ***shift*** of the peak! Tab. B.11 lists the exact numbers.

To check the significance of the movement, the peak position has been fixed to E_{peak}=78.8 GeV. The fits are ***rejected with a chisquare of*** χ^2***/NDF of 33/6*** (maximum value). This corresponds to a probability of $5\ 10^{-8}$ (or 3.8 sigma). This means that the change in spectral index is indeed compatible with the idea of a moving peak, even if we are not able to

$Interval$	T_{obs}	F_0	Peak	χ^2/NDF
in TeV	in h	$cm^{-2}s-1$	GeV	
$[0.5, 1.0]$	21.7	$7.2 \pm 6.1\,10^{-10}$	13.2 ± 5.7	$1.5/4$
$[1.0, 2.5]$	82.9	$2.9 \pm 0.9\,10^{-10}$	30.5 ± 5.7	$2.8/5$
$[2.5, 4.5]$	79.8	$4.0 \pm 0.7\,10^{-10}$	45.4 ± 5.1	$7.3/5$
$[7.0, 10.0]$	23.9	$4.7 \pm 0.7\,10^{-10}$	68.6 ± 7.4	$3.7/5$
$[10.0, 15.0]$	11.8	$5.7 \pm 0.9\,10^{-10}$	78.8 ± 9.25	$2.8/5$

Table B.11: The table shows the fit results of the powerlaw fit with parabolic term. The numbers show that the ***peak of the parabola moves*** from ***13 GeV to 78 GeV*** and that the ***increase*** in flux (in the TeV) might be mainly caused by a ***shift*** of the peak rather than by an increase of the ***peak luminosity.***

$Interval$	T_{obs}	χ^2/NDF
in TeV	in h	Fix curvature a, fix peak at 78.8 GeV
$[0.5, 1.0]$	21.7	$26.2/5$
$[1.0, 2.5]$	82.9	$30.8/6$
$[2.5, 4.5]$	79.8	$33.0/6$
$[7.0, 10.0]$	23.9	$5.4/6$
$[10.0, 15.0]$	11.8	$2.8/6$

Table B.12: To check the significance of the movement, the peak position has been fixed to E_{peak}=78.8 GeV (curve with highest flux) . Three of the fits (curves with lower fluxes) are rejected significantly.

give a good estimate of the real position of the inverse Compton peak. Tab. B.12 lists the chisquare values.

B.2.4 Conclusions and a discussion of the results

Average spectrum of Mkn 421

The spectrum of Mkn 421 shows ***clear evidence*** for a cutoff at $E_c = 3.1 \pm 0.5 \pm 1\,TeV$. The pure power law was rejected at the 6.0 sigma level. The powerlaw has a spectral index of $\alpha = 2.96 \pm 0.04 \pm 0.1$. It was possible to demonstrate that a ***power law with a quadratic curvature term*** of $a = 0.22 \pm 0.05$ describes the shape of the spectrum up to 10 TeV equally well as a power law with exponential cutoff.

Spectrum of the Crab nebula

The fit of a pure power law to the spectrum of the Crab nebula gives a flux constant of $F_0 = 3.0 \pm 0.33 \pm 0.5\,TeV^{-1}cm^{-2}s^{-1}$ and a spectral index of $\alpha = 2.5 \pm 0.1 \pm 0.1$ which is in good agreement with current measurements of other experiments.

Spectra of Mkn 421 during different flare states

The spectra of different fluxes show a ***significant hardening with increasing fluxes***, which is in agreement with the results from the hardness ratio measurement. The decrease of the spectral index with increasing flux was shown by fitting a power law with fixed cutoff (at $E_c = 3.2\,TeV$). A fit with a fixed spectral index was rejected at the 3.7 sigma level.

The fit of a power law with ***fixed spectral index*** and ***free cutoff*** also gives acceptable χ^2-values. The cutoff moves from 1.9 TeV up to 4.3 TeV.

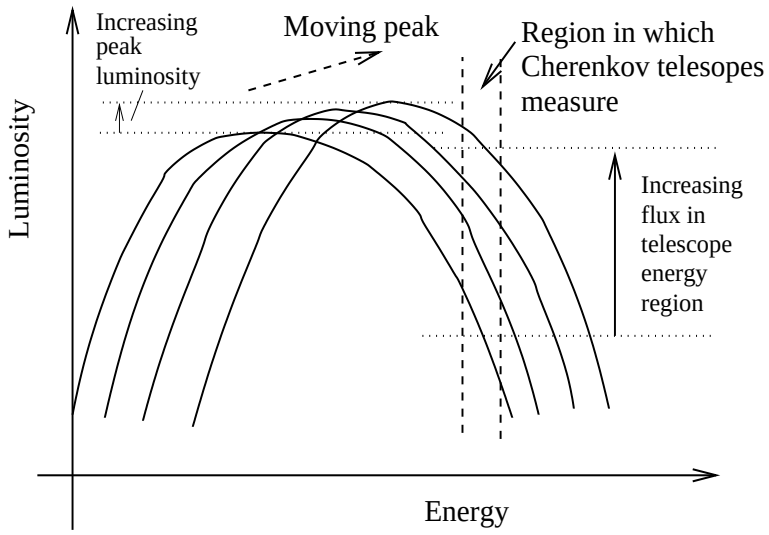


Figure B.29: The figure illustrates the model of the moving peak. The fit of a parabola to the spectra seems to show that a good part of the ***high variability*** of the flux in the TeV range is mainly caused by a ***shift of the inverse Compton peak*** to higher energies (Fig. B.29 illustrates the hypothesis) rather than by ***a high variability of the peak luminosity.***

Mkn 421 flare spectra fit by using a powerlaw with parabolic term

By fitting a power law with a quadratic curvature term (which very roughly approximates an inverse Compton peak) it was possible to show that ***the hardening of the spectrum is compatible with the idea of an moving inverse Compton peak***. The movement of the peak luminosity of the parabola from 13 GeV for low fluxes up to 78 GeV for high fluxes is significant while a ***fixed parabola peak position was rejected with 3.9 sigma***. It must be kept in mind that this rough approximation is not appropriate to estimate the real peak position of the luminosity.

The fit seems to show that a good part of the ***high variability*** of the flux in the TeV range is mainly caused by a ***shift of the inverse Compton peak*** to higher energies (Fig. B.29 illustrates the hypothesis) rather than by ***a high variability of the peak luminosity.***

Conclusion about spectra during different flare states

As a conclusion it can be noted that a ***significant*** change (a hardening for higher fluxes) in the spectrum for different flux levels could be detected. A fixed cutoff and fixed slope fit was rejected as well as a pure power law fit.

But it was not possible to ***differentiate*** between ***different*** models to find out whether the ***cutoff*** changes (fixed spectral index) or rather the ***spectral index*** (fixed cutoff). The hypothesis of a moving peak with a parabolic fit also gives acceptable χ^2-results. The truth might involve a combination of these effects.

Appendix C

Comparison with Mkn 501 and Conclusion

In this final chapter I wish to discuss the measurement results of the large γ-flares of Mkn 421 in 2001 and how the results fit into the global picture of jet emission. I will compare the observational properties of Mkn 421 with the properties of another AGN, Mkn 501, in order to obtain a clearer picture of jet emission. Mkn 501 has become quite famous in TeV-γ-Astronomy. It was the second AGN which emits γ-rays that has been discovered [Qui96] after Mkn 421. It is of virtually the same distance as Mkn 421 and is therefore very interesting for purposes of comparison. It behaves very similar to Mkn 421 but it has longer variability time scales and a different spectrum [Kra01].

This chapter is divided into two parts. In the first part, ***spectral properties*** of the jet emission are discussed. In the second part, ***time scales and correlations*** between x-ray and γ-ray emissions are presented.

C.1 Spectral properties of γ-flares

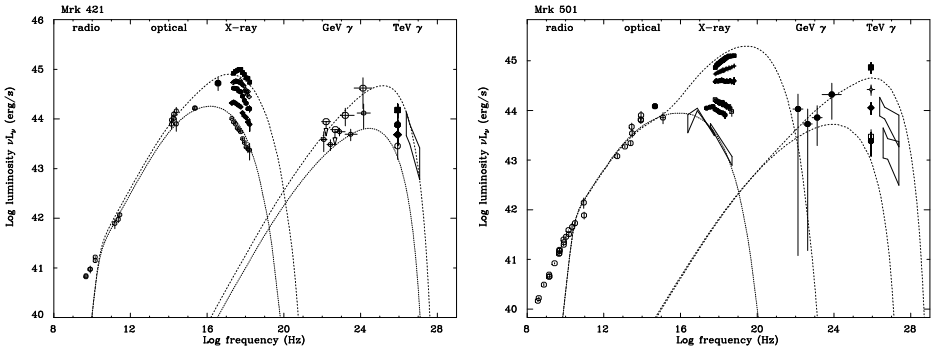
The spectra of Mkn 421 and Mkn 501 exhibit some ***quite similar*** characteristics. Both of them are strongly curved and show a cutoff. On the other hand, the spectrum of Mkn 501 exhibits an ***cutoff at higher energies*** than Mkn 421. The best average (over all flare states) spectrum fit of Mkn 501 is [Kra01]:

$$\frac{dF}{dE} = (10.1 \pm 1.2)\, 10^{-11} \left(\frac{E}{1\,TeV}\right)^{-2.0\pm0.19} e^{-\frac{E}{(6.1\pm1.2)\,TeV}} TeV^{-1} cm^{-2} s^{-1} \quad \text{(C.1)}$$

while for Mkn 421 it is [TW]:

$$\frac{dF}{dE} = (7.0 \pm 0.4)\, 10^{-11} \left(\frac{E}{1\,TeV}\right)^{-1.88\pm0.15} e^{-\frac{E}{(3.1\pm0.5)\,TeV}} TeV^{-1} cm^{-2} s^{-1} \quad \text{(C.2)}$$

Tab. C.1 shows the most important spectral parameters in an overview. Measurements in the x-ray region by ASCA and RXTE [Kat01] show that the synchrotron (luminosity) peak of Mkn 421 changes its ***position*** only slightly ***from 0.5 keV*** in the quiescent state ***to 2.0 keV*** in high flare state (see Figures C.1a) and C.2). For the case of Mkn 501 the position changes ***two orders of magnitude***, from ***1.0 keV*** to ***100 keV*** (see Figures C.1b) and C.2). Fig. C.2 shows a measurement of the synchrotron peak position in the x-ray region as a function of the luminosity.



(a) Mkn 421 (b) Mkn 501

Figure C.1: Plots a) and b) show the spectrum of AGN Mkn 421 and Mkn 501, respectively, over the complete frequency range from radio to TeV for various flare states. The dotted line represents the best fit SSC model for the quiescent/flare state. It can be seen that in the case of Mkn 501 the synchrotron peak shifts two orders of magnitude from 1.0 keV to 100 keV, while for Mkn 421 the position changes only slightly from 0.5 keV in the quiescent state to 2.0 keV in high flare state. Taken from [Kat01].

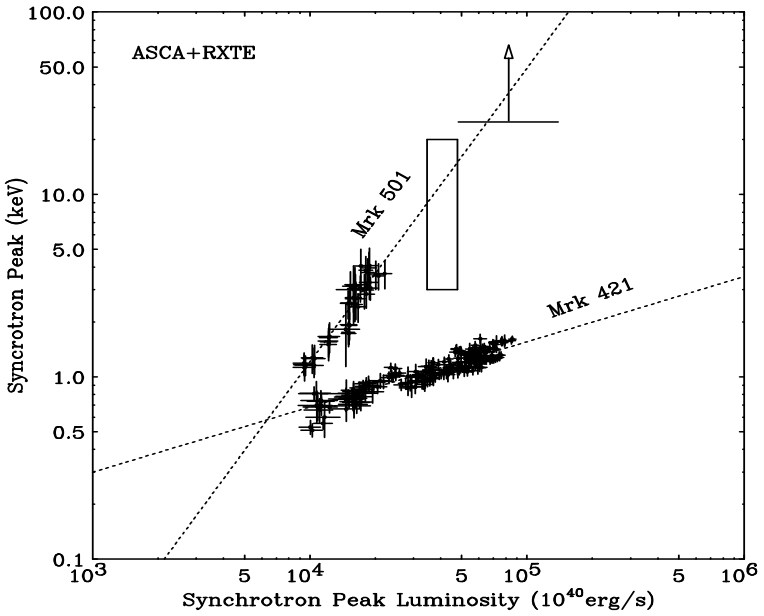
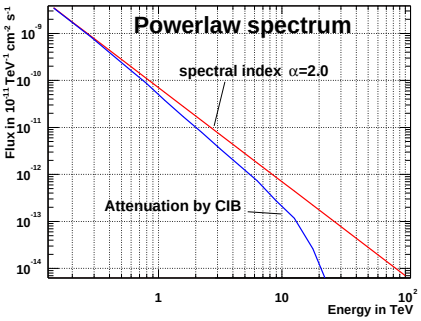
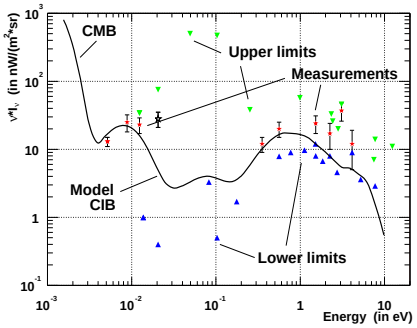


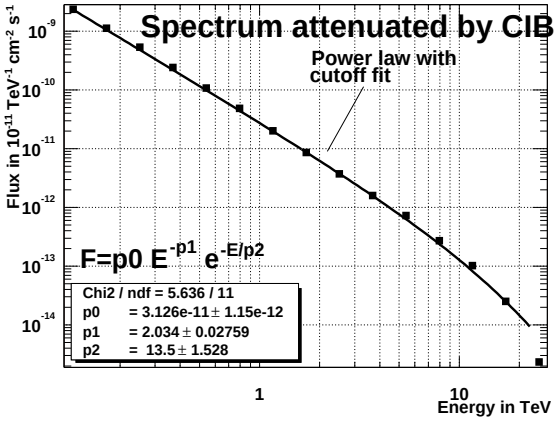
Figure C.2: This figure shows the synchrotron peak position vs. synchrotron peak luminosity for the objects Mkn 421 and Mkn 501 . For the case of Mkn 501 the peak position of the luminosity in relation to the luminosity itself at that point changes much more than for the case of Mkn 421. There the luminosity at peak changes by one order of magnitude while the peak position itself only changes slightly. Taken from [Kat01].



(a) a) Attenuation of high energy gammas



(b) Newst model T. Kneiske and K. Mannheim



(c) Fit to the attenuated spectrum

Figure C.3: Plot a) shows the effect of the cosmic infrared background (CIB) on a power law spectrum with spectral index 2.0 and for a source that is the same distance from earth as Mkn 421. Strong absorption effects that could lead to a cutoff appear mainly above 10 TeV. Plot b) shows the model for the CIB from T. Kneiske and K. Mannheim [Kne02] and is in good agreement with the actual measurements (red points). Green points are upper limits and blue points are lower limits. Plot c) shows a fit to the power law spectrum that has been attenuated by the CIB model. The fit yields an exponential cutoff at 13.5 TeV. The spectrum decreases faster than exponential above 20 TeV which can be seen in the last point which is far below the fit curve.

Parameter	Mkn 421	Mkn 501
Average flux constant F_0	7.0 ± 0.4 [TW]	10.1 ± 1.2 [Kra01]
Spec. index α (with cutoff)	-1.88 ± 0.15[TW]	-2.03 ± 0.19 [Kra01]
Cutoff in TeV	3.1 ± 0.5[TW]	6.0 ± 1.4[Kra01]
Harden. of spec. dur. flare	yes[TW]	yes[Pir01]
Shift of synch. peak in keV	$0.5 - 2.0$ SSC[Kat01]	$1.0 - 100$[Kat01]
Shift of Comp. peak in GeV	$\sim 20 - 200$ [TW], SSC[Kat01]	$10 - 1000$[Kat01, Pir01]

Table C.1: The table lists the spectral properties of Mkn 421 and Mkn 501. For case of Mkn 501 the cutoff is found at higher energies than for Mkn 421. The synchrotron peak of Mkn 501 shifts by two orders of magnitude from low state to high state while the one of Mkn 421 only shifts by a factor of four. The shift of the inverse Compton peak has been guessed by the fit to a SSC model.

This shift in the x-ray range also implies a potential ***shift of the inverse Compton (luminosity) peak*** (in the γ-ray range), which can only be ***estimated*** by a fit to the SSC model [Kat01]. For the case of Mkn 421, it probably shifts from approximately ***20 GeV*** to about ***200 GeV*** while for Mkn 501 it should shift from about ***10 GeV*** up to ***1000 GeV*** [Pir01].

The difference in the average spectrum (i.e. the different ***cutoff***) for the two objects could be explained by the idea that the inverse Compton (luminosity) peak for the case of ***Mkn 501*** shifts to ***higher*** energies than for the case of ***Mkn 421***. A movement of the position of the inverse Compton peak should show up as a ***hardening*** of the spectrum for higher fluxes. This effect has been observed in this work and also by the Whipple telescope [Kre02], which has a much lower energy threshold and by the CAT telescope [Pir01].

Analysis results

As was shown in the last section a ***hardening*** of the spectrum for higher fluxes was observed. Fits with both ***fixed cutoffs and fixed slope*** (-> no change in shape) were ***rejected*** at the 3.8-sigma level, as was a ***pure power*** law fit (rejected with 5.8 sigma). ***Three*** different hypothesis were fitted to the data of five different flux levels. This were

1. a) ***fix-cutoff/free-slope*** hypothesis: By ***fixing the cutoff*** to ***3.3 TeV*** a significant change in the spectral index for increasing fluxes was observed (α decreased from 2.5 to 1.8).

2. b) ***free-cutoff/fix-slope*** hypothesis: By ***fixing the slope*** to α***=2.0*** a significant change of the cutoff was observed (E_c increased from 1.9 TeV to 4.3 TeV).

3. c) A hypothesis of a ***moving*** inverse Compton peak was tested: By fitting a power law with a quadratic term ***(parabolic fit)*** the position of the peak of the parabola shifted significantly from approximately ***15 GeV*** up to ***80 GeV.***

All ***three*** fits gave ***acceptable*** chisquare values.

The parabola is only a ***very rough*** estimate of the true shape of the inverse Compton peak and the values obtained are only demonstrating the compatability with the idea of a moving peak and say nothing about the ***true*** inverse Compton (luminosity) position. The fit of the parabola ***suggests*** that the ***high variability*** of the flux in the region between 500 GeV and 20 TeV (where Cherenkov telescope measure) might ***not only originate*** from a variablity of the (inverse Compton) peak luminosity but also from the ***shift*** of the inverse Compton peak, since the slope of the spectrum in the region between 500 GeV and 20 TeV is ***very steep*** and any movement of the peak immediately results in a ***large change*** in the flux there (See Fig. B.29).

It was ***not possible*** to clearly ***accept*** or ***reject*** any one of these hypothesis. The true behaviour might involve a combination of these effects.

Parameter	Mkn 421	Mkn 501
Correlation x-ray/TeV (1 day bin)	0.74 ± 0.12[TW]	0.76 ± 0.14[Kra01]
$Flux^{b}_{xray} \propto Flux_{\gamma}$: b	1.0 ± 0.2[TW]	1.96 ± 0.07 [Kat01]
Doubling time of flare (TeV)	15 min.[TW]	$\sim 2\,h$[Kra01]
Doubling time of flare (keV)	$\sim 15\,min$[Car99]	$\sim 2\,h$[Kra01]
Average flare duration (TeV)	$1-3\,h$[TW]	$\sim 2\,days$[Kra01]

Table C.2: The table lists the most important properties concerning time scales and correlations for the objects Mkn 421 and Mkn 501.

The question about the cutoff in the spectra of Mkn 421 and Mkn 501

In an attempt to clarify the ***origin of the cutoff***, one could start examining the effects of the CIB (Fig. C.3). The question is whether the cutoff is an ***intrinsic*** feature of the real emission spectrum or a result of the ***absorption*** of γ-rays due to the cosmic infrared background (gamma ray horizon). In Fig. C.3a), the ***effect*** of the CIB absorption on a power law spectrum of α=2.0 is displayed. The model used for the CIB is [Kne02] that fits the actual CIB measurements very well. Due to the ***steepness*** of the spectrum, a true ***cutoff effect*** can only be observed ***above 10 TeV.*** A fit to the attenuated power law spectrum yields an exponential ***cutoff at 13.5 TeV*** while it has to be kept in mind that the attenuation above 10 TeV is much stronger than exponential. This suggests that the ***dominant*** reason of observed cutoff in the spectra of both Mkn 421 and also of Mkn 501 might favor the hypothesis of an intrinsic cutoff (***increasing cutoff energy with increasing flux***) and a ***moving inverse Compton peak***.

C.1.1 Time scales, correlations and jet models

Time scales of flares and correlations between x-ray emission and γ-ray emission also provide us with hints of the physical mechanisms inside the jets. Tab. C.2 provides an overview of the most important parameters.

Correlation of x-ray emission and γ-ray emission

Both sources show a ***strong correlation*** between the x-ray flux and the γ-ray flux. For Mkn 421, a correlation coefficient of 0.74 ± 0.12 (this analysis, one day bins) and for Mkn 501, a correlation coefficient of 0.76 ± 0.12 [Kra01] was observed. This correlation provides hints about the mechanism of γ-ray production. Today it is widely assumed that high energy gammas are produced via inverse Compton up-scattering of soft photons. The question is whether the soft photons emerge from the ***accretion disc*** or whether they are produced ***inside*** the jet. The high correlation can only be explained if both production mechanisms are ***coupled*** and very close to each other in distance. If the soft photons (i.e. x-ray photons) would originate from the accretion disc then the correlation would be ***zero*** or even ***negative*** (x-ray dip). This means that an existing correlation clearly ***favors*** a SSC model and as a consequence, potential quasi-periodicities (like the 23 +- 2 day periodicity of Mkn 501 [Kra01]) ***cannot*** be explained by a ***modulation*** of soft photons ***emerging from the accretion disc*** [Bed96]. Quasi-periodicities should be explained by other models like helical jets, periodic ejection of the jet by a big rotating object close to the innermost stable orbit of the (Kerr-) black hole [Mas99] or binary black holes [Beg80, Man00].

The ASCA x-ray measurement

Interesting results are obtained from the ASCA-x-ray/γ-ray correlation study that also shows a ***strong correlation*** between x-ray flux (ASCA measurement from April 1998) and the γ-ray flux. In addition, we divided the x-ray flux into a ***slow flare*** component

(of several days duration) and a ***fast component*** (of a few hour duration). It appears that the correlation between x-ray and γ-ray ***increases*** from 0.63 +- 0.18 to 0.70 +- 0.16 when the slow x-ray component (the background) is subtracted. However, the increase is not significant and does not allow further conclusions.

Relation between the x-ray flux and γ-ray flux

In the case of ***Mkn 501*** the x-ray/γ-ray relation is ***quadratic*** ($Flux_{xray} \propto Flux_{Gamma}^{1.96\pm0.07}$ [Aha99/2]) while for ***Mkn 421*** it is ***linear*** ($Flux_{xray} \propto Flux_{Gamma}^{1.0\pm0.2}$) (see the previous chapter). The latter measurement is in agreement with [Aha99/1] and is also confirmed by the ***ASCA-x-ray/γ-ray*** correlation study (see the previous chapter). This difference between the objects can be explained by the fact that the peak position (of the inverse Compton peak) of Mkn 501 shifts ***much stronger*** with changes in its luminosity.

The argument is the following. The change in flux that we observe in the small TeV window might not only originate from a ***real increase*** in the luminosity (at peak) but also from a ***movement*** of the inverse Compton peak. Since the movement for the case Mkn 501 is presumed to be ***larger***, the change in ***flux*** in the 500 GeV to 20 TeV window is expected ***much larger*** (it is ***quadratic***) in contrast to Mkn 421 in which case the change in ***flux*** (in the TeV window) should be ***less*** (since the inverse Compton peak is ***moving less*** it is ***linear***). This means that the difference in the relation between x-ray fluxes and γ-ray fluxes is strongly related to the ***question*** at which ***positions*** we find the synchrotron/inverse Compton peaks, which ***shape*** they have and in which ***energy intervals*** we measure the flux.

In the future, more precise multiwavelength measurements will provide the possibility to reject or accept different SSC model theories.

Observable and intrinsic parameters of the classic SSC model

The simplest SSC model (assuming only a single ***spherical emission region***) was presented in the introduction chapter (in detail in Appendix A) and has the following seven independent parameters: The normalization constant K, the energy break energy γ_b, the radius of the emission region R, the magnetic field B, the Doppler factor δ and the positions of the peak luminosities of the synchrotron radiation ν_s and of the inverse Compton radiation ν_C [Ino96].

The spectral observations in the x-ray and the γ-ray energy region are the ***shape*** of the spectra, the two peak ***luminosities***, the ***variability time scales*** and a potential ***time lag*** between the x-rays and the γ-rays. The structure of the equations allows to express strong constraints on B and δ [Tav98]. For the case of Mkn 421, using the single emission region model and a variability with a ***very conservative*** time scale of $t_{var} = 1\,h$, the magnetic field B and the Doppler factor δ result in $B \simeq 0.25$ Gauss and $\delta \simeq 25$. This ***atypical high*** Lorentz factor is meant to provide an explanation for the ***very rapid*** variability time scales observed.

Variability t_{var} and the size of the emission region R

The time scales of the flares are quite different for the two sources. Mkn 421 shows ***doubling times*** of equal to or faster than ***15 minutes*** (see the analysis of the previous chapter) while Mkn 501 has doubling times of about 2 hours [Kra01]. The typical (fast) flare duration (FWHM) for Mkn 421 is approximately 1-3 hours [TW] while for Mkn 501 it is about 2 days .

These variabilities can be explained by varying the ***dimensions*** of the emission regions of the jets of these objects and by ***different Doppler factors*** due to varying bulk jet ***Lorentz factors***.

For the case of Mkn 421, for reasons of causality, the small variability time scale of $t_{var} = 15$ minutes puts ***hard constraints*** on the ***size*** of the emission region and/or on the Doppler factor δ:

$$R \leq c t_{var} \delta \tag{C.3}$$

This implied that for Mkn 421, assuming a ***typical*** Lorentz factor of $\delta = 10$ (which cannot be much higher for Blazars as explained in the introduction chapter), that the emission region must be ***very small***, less than $R = 2\,10^{12}\,m \simeq 10\,AU \simeq 10^{-4}\,pc$. Typical jet scales are approximately $0.01 - 0.03\,pc$. Jet dimensions are ***strongly constrained*** by the ***transparency or photon-photon opacity condition*** which sets lower limits [Mas99] on them. Therefore jet scales with sizes of order $10^{-4}\,pc$ are ***virtually ruled out*** for the observed ***luminosity***.

The ***other possibility*** for this model is to ***increase*** the Lorentz factor up to values of δ=80. This scenario is ***unrealistic***, especially since it is known that the (cooled) radio blobs (the superluminal motion is only 2-3) of these objects are ***slower*** than those of typical blazars and radio galaxies.

Shock-in-jet models

Obviously the standard jet model of a spherical emission region ***seriously breaks down*** when it attempts to explain the ***very short variability time scales***. For this reason, new models have been developed which are able to explain the fast variabilities ***without*** increasing the Doppler factor. All of them introduce ***moving laminar shock fronts*** inside the jet instead of a spherical emission region. Thin ***'sheets'*** of shock fronts move inside the jet towards the observer. Typical source dimensions have ***radii*** of approximately $R = 10^{14}\,m \sim 0.01\,pc$ and ***thicknesses*** of $d = 7 \cdot 10^{11}\,m \sim 5 \cdot 10^{-5}\,pc$. These so-called ***shock-in-jet*** models provide a natural explanation for very short variabilities [Mas99] (a more detailed discussion can be found in [Sal98]).

When looking at the recorded flares of Mkn 421 (intra day variability, lightcurves for each night) one gets the impression that a ***large flare*** (several days duration) is mainly a ***superposition*** of many ***small flares.*** This image was used as a ***basis*** for the flare model that was fitted to the intraday lightcurves (last chapter, fitting of flare times): The flux of a single flare is $F(t) = a + \frac{b}{\left(2^{(t-t_0)/c} + 2^{-(t-t_0)/d}\right)}$. In this model a ***constant background*** 'a' (the pile-up of many flares), an ***exponential rise time*** 'c' (the acceleration inside laminar shocks-in-jet) and an ***exponential cooling time*** 'd' (the cooling via synchrotron radiation and inverse Compton scattering) was assumed. This picture naturally emerges from shock-in-jet models because there one expects many shocks in ***parallel*** which cross the jet at the same time.

In addition, the measurements of ***ASCA*** in the x-ray energy region (last chapter, correlation measurements between x-rays and γ-rays) seem to show a ***slow flare*** component and ***fast overlapping*** flares. The slow component covering several (about seven) days refers to the ***emerging of a blob*** and the fast short flares refer to the ***shocks-in-jet*** that traverse the blob. This would explain why the correlation seems to increase if the slow component (in the x-ray) is subtracted from the total before calculating the correlation.

Potential timelag of the x-rays

In the analysis of the Mkn 421 flares in 2001 we have observed a hint that the x-rays may be ***delayed*** with respect to γ-rays by approximately ***10 hours.***

The ***observed timelag*** of the x-rays had only an estimated significance of ***three sigma*** since the mentioned systematical error is large. However, a potential timelag can originate from four possible situations:

1. ***Asymmetric x-ray flare shape:*** A significantly ***shorter rise time*** than ***fall time*** of the x-ray flares could result in an effect as seen in the last chapter. The resulting shape of the discrete correlation function would also be ***asymmetric*** [Bad98, Jor01] and the observed effect thus cannot be interpreted as a ***physical*** timelag. This means that a non-linear and complicated ***transfer function*** of the ***x-ray*** flux to γ***-flux*** could result in an ***asymmetry*** of the discrete correlation function around zero time lag. Unfortunately, the ASM data does not allow to give answers about the shape of x-rays flares and the transfer function.

2. ***Opacity effects:*** Different ***optical opacities*** of the jet for x-rays and γ-rays can have the effect that x-rays arrive later. The x-rays are ***trapped*** and scattered inside the jet and can only escape the jet when it is sufficiently cooled down whereas the γ's ***escape immediately*** [MagCom].

3. ***Cooling effects:*** Another explanation could involve ***cooling*** of the high energy electrons. The $e^{\pm}$in the jet are cooled down via synchrotron radiation and inverse Compton scattering. A natural consequence of electron cooling is a ***time lag*** of soft x-rays with respect to hard x-rays. The argument is the following: In high state the x-ray spectrum is ***hard*** and becomes ***softer*** as it cools down. The more energetic x-rays ***appear first*** and later (after cooling) ***soft x-rays dominate*** the spectrum. This introduces the time lag. Since the inverse Compton peak is sort of ***mirror*** of the synchrotron peak, a time lag between soft x-ray and γ-rays is also observed. The estimated order of magnitude of the time lag between soft and hard x-rays is (i.e. the decay time of electron energy) is given by [Dar97]:

$$t_{lag} \sim 10^3 \frac{H_{\perp}^{-3/2}}{Gauss} \frac{E_{\gamma}^{-1/2}}{keV} \delta^{-1/2} \sim \ a\, few\, hours \tag{C.4}$$

   H is the magnetic field in Gauss, E_{γ} id the energy of the electron in keV and δ is the Lorentz factor. ***Common*** electronic jet models [Dar97](with a spherical emission region) have ***difficulties*** to explain the fast cooling times of electrons that are at the same time ***efficiently accelerated*** to ultra high (TeV ranges) energies. Shock-in-jet models ***avoid*** these problems because acceleration and cooling are separated and happen at ***different*** positions in the jet [Sal98]. It should be mentioned that short fall times that have been measured for γ-ray flares don't imply that x-ray flares also have short fall times (cooling). It is probable that x-ray flares have a different shape than γ-flares.

4. ***Quantum gravity:*** The timelag could be a ***first sign of quantum gravity*** effects where theoreticians expect changes in the ***speed of light*** for photons with energies close to the Planck mass. The time lag of x-rays would point to an ***increase*** of the speed of light for increasing photon energies [Ame00, Ame96]. Even that the energy of TeV photons is far away from the Planck scale, the long distance from Mkn 421 to earth could help to amplify even tiny effects.

As a conclusion we can only give an upper limit on the time lag of about 13 ± 3 hours.

Future multiwavelength observations with advanced instruments will clarify these questions in greater detail. Larger, future Cherenkov telescopes record with higher statistics and will provide the possibility to find time lags within GeV/TeV data only.

Fast variability and the mass of the black hole

The variability time scale is approximately 15 to 20 minutes. The appearance of the fast flares is ***entirely random***. Therefore, they are not a result of helical jets or objects close to the innermost stable orbit of the black hole but rather an effect of the jet dynamics itself. A very plausible explanation for this are the ***laminar shock fronts*** inside the jets, which was

just discussed. The jet has ***not a direct*** relation to the size of the central black hole. This implies that an estimate of the mass of the black hole ***cannot*** be found here.

Generally, it is assumed that the nucleus of Mkn 501 has a mass of 10^8 solar masses. Mkn 421 is considered to be smaller than Mkn 501 by 2 orders of magnitude, with an estimated mass of 5×10^6 solar masses because of the shorter variability. The mass of the black hole is generally presumed to scale with $m \propto t_{var}$ because the circling time around the black hole scales linearly with mass.

C.2 Outlook

Future Cherenkov telescopes as MAGIC on the island of La Palma, HESS in Namibia or VERITAS have ***lower energy thresholds***, ***higher sensitivities*** and much ***higher effective detection areas.*** They record with a much higher event rate and achieve ***better statistics*** and a ***better separability*** of γs and hadrons. It will be possible to clarify to a higher level the following questions which have been touched in this work:

1. It might be possible to directly observe the ***energy of the peak luminosity*** with lower energy thresholds or at least to give precise estimates by fitting the SSC model. It will be for sure possible to significantly clarify the question whether ***peak moves*** with varying luminosity or not. This gives an answer to the question of the ***origin of the cutoff***.

2. Preciser data might allow to see if the observed spectrum contains ***contributions of gamma photons from π^0-decay*** and will therefore detect ***hadronic components*** in the jet and confirm whether the jet is purely electronic or not. Neutrino experiments like Amanda and Ice Cube at the South Pole and Antares in the Mediteranian Sea will help answering the question about hadronic contributions in jets (the detection of neutrinos prove the existence of hadrons). The detection of hadrons would help to solve one of the most urgent questions in cosmic ray physics: Where do the high energy hadrons come from ?

3. Since the threshold of the new instruments is ***lower, more objects*** behind the actual gamma ray horizon will be seen. This gives information about the ***exact position*** of the gamma ray horizon and therefore about the ***evolution of star formation*** during the history of the universe (cosmic infrared background) and also about cosmic parameters as for example the ***cosmological constant*** [Bla01]. The determination of the position of the gamma ray horizon will also clarify the ***origin of the cutoff*** in the spectra.

4. Measurements over two to three orders of magnitude of the high energy scale (30 GeV to 30 TeV) with high statistics will open the possibility to detect ***possible time lags*** of fluxes in lower energy regions (<100 GeV) to fluxes in higher energy regions (>1 TeV) ***within the same γ-ray dataset*** only.

5. With higher statistics ***rise times*** (acceleration) and ***fall times*** (cooling) of the flares can be measured in a more accurate way. It will be possible to observe whether long flares are a ***superposition*** of ***many small flares*** or if other components come into play. This will gives us a better understanding of the jet model and the mechanisms that are active in jets.

Simultaneous multiwavelength campaigns with precise x-ray measurements are mandatory for the following reasons:

1. It will be possible to determine the exact ***transfer function*** of x-ray flux to γ-flux, the acceleration time and the cooling time of the high energy particles in the jet.

This will tell us which processes take place in jets, whether there are ***laminal shock fronts*** or not and how γ-rays are produced.

2. When the transfer function is known, the question about the ***time lag*** of x-rays to γ-rays will be solved in a more accurate way. It might be possible to detect signals of quantum gravitation.

Appendix D

Summary

Today astroparticle Physics is one of the most booming fields in modern physics. Gamma ray astronomy is a ***major tool*** to analyze and understand the ***origin*** of cosmic rays. The BL LAC object, ***Mkn 421***, belongs to a class of objects that ***emits*** cosmic rays, the so-called ***active galactic nuclei*** (AGN). In a stochastic but regular manner it switches from a ***quiescent*** state to a state of ***high activity*** in which it starts to emit photons in the x-ray and TeV energy range. From ***February 2001 until April 2001*** it exhibited especially large and long ***flares,*** which have been observed by many telescopes in the TeV energy range as well as by the CT1 telescope of the HEGRA collaboration at La Palma and in the x-ray energy range by the ***All Sky Monitor*** (ASM) of the RXTE satellite. Another flare happened in April ***1998*** which has been observed by the ***ASCA*** x-ray satellite and simultaneously by ***several*** Cherenkov telescopes including ***Wipple***. I present here an analysis of the la***rge gamma flares of 2001*** and also of the short flare in April 1998.

The ***classical*** analysis and gamma/hadron separation methods of Cherenkov telescopes have been improved by several ***new techniques***, which have been tried and tested for the first time here. A complete analysis package of approximately 50.000 lines of code has been developed in ***ROOT/C++*** which contains, apart from the items of a standard analysis for Cherenkov telescopes, the following ***new features:***

- A ***modified linear discriminant analysis*** to enhance the ***gamma/hadron separation***. It is also used as a tool to ***quantify*** the separation capability of different ***sets*** of (new) image parameters and new image cleaning algorithms that have been tested and studied in detail.

- Different algorithms that perform so-called ***image cleaning,*** a technique which is widely used to remove noise background in the image, have been tried and tested with respect to ***improvement*** of ***gamma/hadron separation.***

- Introduction of ***new*** image parameters that improve ***gamma/hadron separation*** and the

- Introduction of ***weights*** in the calculation of the ***image parameters***, the so-called ***'Hillas'*** parameters, which improve ***gamma/hadron separation***

- A systematic algorithm that tests virtually all possible combinations of parameters and image cleaning to find the one with optimal ***gamma/hadron separation.***

- A new method to correct the ***mispointing*** of the telescope. This ensures zenith angle independent integrated flux measurements.

- Unlike the ***usual*** case, the characteristic parameters and features of the emission spectrum have not only been deduced from the ***unfolded flux spectrum*** but rather,

vice-versa, by fitting ***raw excess event energy distribution.*** In this way an unsatisfactory and potentially unreliable ***unfolding process*** is avoided and the fit results (i.e. the spectrum parameters) have less errors and are more robust.

In addition to these improvements, the ***standard items*** have also been implemented. They are:

- A program to ***estimate the energy*** on the basis of the ***least squares*** method. It achieves an average energy resolution of approximately 24 %.
- A program to calculate ***effective areas.***
- A program to calculate the ***spectrum*** by ***unfolding*** the ***energy resolution*** obtained from MC generators. This not only solves the ***spill-over*** problem from higher energy bins to lower energy bins, but it also corrects for systematic biases introduced by an energy estimate that is ***not*** completely ***linear.***
- A program to calculate the time dependent integrated flux (***light curve***) in various energy ranges. The According effective areas have been determined from the actual shape of the spectrum, which was previously calculated.
- A program to calculate the ***hardness ratio*** of the lightcurve.
- A program to calculate the ***correlation*** between x-ray lightcurves and γ-ray lightcurves.
- A program that ***simulates the night sky background*** in the MC gamma data sample.

The improved analysis has been ***applied*** to the measurements of the ***large gamma flares*** in 2001 of Mkn 421. 249 hours of observation time have been accumulated, which allows several detailed studies with good statistics:

- A detailed lightcurve with 20 minute bin size granularity, which reveals the extremely ***fast flares*** of this object. The minimal doubling and halving times and the typical FWHM duration of fast flares were estimated.
- The flux dependence of the ***hardness ratio*** was examined for different energy intervals. Evidence of a change in spectral index below 2 TeV from low to high fluxes has been observed.
- The ***correlation*** between the γ-TeV-flux and the x-ray light flux as measured by the ASM/RXTE (0.5 keV to 10 keV) has been calculated.
- The ***average spectrum*** and it characteristics have been calculated and determined.
- The ***spectral behaviour*** of the source during ***different flare states*** has been carefully analyzed in order to see if the effect observed, using the hardness ratio, could also have been observed directly from the spectrum.

In brief, the ***final results*** are the following:

- The spectrum of the Crab nebula was calculated and yielded a pure powerlaw with a spectral index of $\alpha = 2.5 \pm 0.1 \pm 0.1$ and a flux constant of $F_0 = 3.0 \pm 0.33 \pm 0.5\ TeV^{-1}cm^{-2}s^{-1}$ which is in good agreement with the measurements of other Cherenkov telescopes.
- A pure power law fit to the averaged spectrum of Mkn 421 was ***rejected*** by a chisquare of χ^2/NDF=74/9. A ***power law fit with exponential cutoff*** was accepted by a chisquare of χ^2/NDF=6.3/8. The spectral index is $\alpha = 1.88 \pm 0.15 \pm 0.1$. The cutoff was found at $E_c = 3.1 \pm 0.5 \pm 1.0$ TeV.

- The spectrum for ***five different flux levels*** were calculated and it was possible to show that the spectrum changes significantly with increasing flux: It gets ***harder.*** The hardening of the spectrum in the region below 2 TeV could be ***verified*** by an independent analysis using the hardness ratio.
- ***Three*** different hypothesis were fitted to the five spectra. For statistical reasons only functions with ***two free*** parameters were used. All three hypothesis yielded ***acceptable*** chisquare values. The hypothesis were:
  - A power law hypothesis with ***fixed cutoff*** at $E_c = 3.2\,TeV$ and free slope: The spectral index ***decreased*** significantly with increasing flux (a fixed spectral index with fixed cutoff was rejected at the 3.8 sigma level).
  - A power law hypothesis with ***fixed slope*** $\alpha = 2.0$ and free cutoff: The cutoff ***moved*** significantly from 1.9 TeV up to 4.3 TeV.
  - A power law with a ***parabolic (quadratic)*** term: The peak (luminosity) of the parabola ***moved*** significantly from 13 TeV up to 78 TeV (a fixed parabola peak was rejected at the 3.5 sigma level) indicating that the change of the spectrum is ***compatible*** with a ***moving*** inverse Compton peak. The peak of the parabola is only a ***very rough estimate*** of the real luminosity peak. Still, the ***movement*** was ***significant.***
- The TeV lightcurve has been examined at ***single night*** basis. The rise and fall time were estimated in ***three*** independent different manners.
  - A fit of a simple flare model allowed to determine some flare parameters as ***rise time***, ***fall time*** and ***average duration*** (FWHM). The model assumes a ***constant background*** and ***exponential*** rise and fall times. The ***average short flare duration*** (FWHM) is of the order of ***1 to 3 hours.*** The model shows that most of the fast flares have ***doubling rise times*** and ***halving fall times*** of about ***25 minutes***. No significant ***difference*** between rise and fall times could be found.
  - The doubling rise time and halving fall time were calculated by using a ***4-point/three-straight-line*** fit. The fit values show that the fastest flares have doubling times of ***at most 15 to 25 minutes.***
  - The doubling rise time and halving fall time were calculated by only measuring the rise and fall between ***two points.*** Since this method is ***sensitive*** to fluctuations in the flux data points and therefore can result in ***underestimated*** rise/fall times, the slope of the straight line connection between the two points was calculated by ***subtracting/adding*** one sigma on each beforehand (***worst case scenario***). The fastest flares showed rise and fall times of ***at most 15 minutes.***
- The x-ray flux measured with ASM/RXTE shows a ***clear and significant correlation*** to the TeV data of approximately ***0.74+-0.12*** (for one day bins).
- Both correlation studies, with CT1/ASM data and ASCA/TeV data, yielded a ***linear relation*** between x-ray flux and γ-ray flux.
- A hint of a ***potential time lag*** (about 10 h+-2 h) of the x-rays to the γ-rays showed up. It appears that the hypothesis of a x-ray time lag of several hours compared to γ-rays could ***not*** be ***significantly*** proven because of the uncertainty in the ***estimation of the systematic error***. However, every correlation curve shows a systematic time lag for the complete range of time-bin widths, from 1 day down to 0.5 h. The apparent time lag could be a result of an ***asymmetric shape*** of the x-ray flares with a short rise time and a very long fall time. Therefore only an ***upper limit*** on the time lag of 13 ± 3 hours is given here.

- A flare of Mkn 421 in April 1998 as measured by the ASCA x-ray satellite was analyzed and the correlation to γ-ray measurements of several telescopes was calculated. A ***significant correlation*** was found. The x-ray data was splitted into a ***slow flare*** component (seven days duration) and a ***fast*** flare component (several hour duration). The correlation of the TeV flux to the total x-ray flux was ***0.63+-0.18*** and increased to ***0.70+-0.16*** when the ***slow*** component was ***subtracted*** from the lightcurve. The increase is not significant enough to give further conclusions.

Appendix E

Appendix A: Theory and calculations

E.1 Verification of the NSB distribution function

In the introduction chapter the ***excess noise factor F*** has been introduced with the following general definition:

$$F^2 = \frac{Signal^2_{input} / \left(\sigma^2_{input} - \sigma^2_{el}\right)}{Signal^2_{output} / \left(\sigma^2_{output} - \sigma^2_{el}\right)} \tag{E.1}$$

The noise at the input and at the output is understood as noise ***without*** electronic noise from the amplifier $\sigma^2_{noise} = \sigma^2_{real} - \sigma^2_{el}$.

The output distribution of a PMT (Gain normalized to Gain=1) when exposed to diffuse night sky background (NSB) has been described as the ***Poisson distributed sum of Gaussian normal distributions***.

$$f_\lambda(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\,\lambda^n}{n!} \frac{e^{-\frac{(x-n)^2}{2\sigma_n^2}}}{\sigma_n\sqrt{2\pi}} \tag{E.2}$$

with a ***variance*** of

$$\sigma_n^2 = \quad n\left(F^2 - 1\right) + \sigma^2_{el} \tag{E.3}$$

λ is the ***average*** amount of photoelectrons that hit the first dynode, ***n*** is the number of photoelectrons, σ_n^2 is the variance of the individual photoelectron peak and σ_0^2 is the ***variance of the pedestal*** which is equal to the ***electronic noise*** contribution $\sigma_0^2 = \sigma^2_{el}$.

Now we want to ***cross check*** if the proposed output distribution (Equ. E.2) is ***consistent*** with the definition of the excess noise factor.

1) We calculate the ***average*** of $f_\lambda\,(x)$:

$$\langle f_\lambda\,(x)\rangle \quad = \quad \int_{-\infty}^{+\infty} f_\lambda\,(x)\; x\,dx \tag{E.4}$$

$$= \quad \int_{-\infty}^{+\infty} \left(\sum_{n=0}^{\infty} \frac{e^{-\lambda}\,\lambda^n}{n!}\, x\, \frac{e^{-\frac{(x-n)^2}{2\sigma_n^2}}}{\sigma_n\sqrt{2\pi}} \right) \tag{E.5}$$

$$= \quad \sum_{n=0}^{\infty} \frac{e^{-\lambda}\,\lambda^n}{n!}\, n$$

$$= \quad \lambda \;\; OK.$$

2) We calculate the ***variance*** of $f_\lambda(x)$:

$$\begin{aligned} var_x(f_\lambda(x)) &= \left\langle (x-\overline{x})^2 \right\rangle \\ &= \int_{-\infty}^{+\infty} \left(\sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} (x-\lambda)^2 \frac{e^{-\frac{(x-n)^2}{2\sigma_n^2}}}{\sigma_n\sqrt{2\pi}} \right) \end{aligned} \tag{E.6}$$

With $(x-\lambda)^2 = (x-n)^2 + 2(n-\lambda)(x-n) + (n-\lambda)^2$ we get

$$\begin{aligned} var_x(f_\lambda(x)) &= \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left\{ \sigma_n^2 + (n-\lambda)^2 \right\} \\ &= \lambda\left(F^2-1\right) + \sigma_{el}^2 + \lambda^2 - 2\lambda^2 + \lambda^2 + \lambda \\ &= \lambda F^2 + \sigma_{el}^2 \end{aligned} \tag{E.7}$$

3) We introduce $Signal_{input} = Signal_{output} = \lambda$, $\sigma_{input}^2 = \lambda$ and $\sigma_{output}^2 = \lambda F^2 + \sigma_{el}^2$ in Equ. E.1:

$$\frac{\left(\lambda/\sqrt{\lambda}\right)^2}{\left(\lambda/\sqrt{\lambda F^2}\right)^2} = F^2 \;\; OK. \tag{E.8}$$

E.2 The synchrotron self Compton (SSC) model

In the SSC model[1] the gamma-ray photons are produced by the inverse Compton scattering of soft photons. These soft photons are generated by the same electrons stemming from the synchrotron emission (see Fig 1.6). The model discussed here has been described in [Ino96, Blo96]. It is based on the following assumptions:

1. A ***spherical emission*** zone which is stationary and has a parameter R that describes all length scales.
2. An ***electron distribution*** $N(\gamma)$ which is parameterized as a ***broken power law*** (K is a normalization constant, α is the spectral index of the shock acceleration (eg. $\alpha = 2.2$), γ is the electron lorentz factor and γ_b is the breaking energy)

$$N(\gamma) = K\gamma^{-\alpha}\left(1 + \frac{\gamma}{\gamma_b}\right)^{-1} \tag{E.9}$$

3. The soft photons[2] needed for the inverse Compton scattering are ***synchrotron photons***.
4. All particle and photon distributions are ***isotropic*** in the jet frame.

The electron spectrum has the shape of a ***broken*** power law because of ***cooling effects*** during the shock acceleration process. The high energy electrons are continously cooled by synchrotron radiation and inverse Compton scattering. The breaking energy results to be

$$\gamma_b = \frac{3mc^2}{4\left(u_b + u_{soft}\right)\sigma_T R} \tag{E.10}$$

with magnetic field density

$$u_B = \frac{B^2}{8\pi} \tag{E.11}$$

The electron energy is given in terms of the Lorentz-factor γ, B is the magnetic field in Gauss, K is a normalization constant, σ_T is the Thompson cross section and m is the mass of the electron in EGS units. γ_b breaks the spectrum by one power in the index.

Synchrotron radiation

The ***basic formulas*** for ***the synchrotron radiation*** are described in the following.

1. ***The spectral density*** is
$$I_\nu = \frac{j_\nu}{\alpha_\nu}\left(1 - e^{-\alpha_\nu R}\right) \tag{E.12}$$
2. where j_ν and α_ν are the ***emission*** and ***absorption coefficient*** for the synchrotron radiation, which are calculated in the following way:
$$j_\nu = c_2 B \int_{\gamma_{min}}^{\gamma_{max}} d\gamma N\left(\gamma\right) F\left(\frac{\nu}{c_1 B\gamma^2}\right) \tag{E.13}$$
and
$$\alpha_\nu = -c_3 B \frac{1}{\gamma^2}\int_{\gamma_{min}}^{\gamma_{max}} d\gamma\gamma^2 \frac{\partial}{\partial\gamma}\left[\frac{N\left(\gamma\right)}{\gamma^2}\right] F\left(\frac{\nu}{c_1 B\gamma^2}\right) \tag{E.14}$$

[1] Note: The two expressions 'hard photons' and 'high energy γ-rays' denote the same objects.
[2] Note: In general soft photons are (soft) x-ray photons or (hard) UV photons

with

$$F(x) = x \int_x^\infty K_{\frac{5}{3}}(x')dx' \tag{E.15}$$

$K_{\frac{5}{3}}(x)$ is the incomplete Bessel function of fractional order $\frac{5}{3}$and the constants are $c_1 = \frac{3e}{4\pi mc}$, $c_2 = \frac{\sqrt{3}e^3}{8\pi^2 mc^2}$ and $c_3 = \frac{\sqrt{3}e^3}{8\pi m^2 c^2}$. The electron density $N(\gamma)$ was shown in Equ. E.9.

Inverse Compton scattering

The emission coefficient of the inverse Compton scattering is obtained from the energy spectrum of the electrons Equ. E.9 as follows [Ino96]:

1. The ***emission coefficient***

$$j_\nu = \frac{h}{4\pi}\varepsilon q(\varepsilon) \tag{E.16}$$

and

$$\nu = \frac{m_e c^2}{h}\varepsilon \tag{E.17}$$

can be calculated from the

2. ***Differential hard photon production rate*** shown in Equ. E.18 (number of hard photons ε produced per energy interval per unit volume per unit time) which is a convolution of the electron distribution and the soft photon distribution (where the photon density is $n(\varepsilon_0)$, the soft photon energy is ε_0, the final photon energy is ε and the electron energy is γ , and everything is in units of electron masses)

$$q(\varepsilon) = \int d\varepsilon_0 n(\varepsilon_0) \int d\gamma N(\gamma)\, C(\varepsilon\,\gamma\,\varepsilon_0) \tag{E.18}$$

The probability for a soft photon to be up-scattered is called the

3. ***Compton kernel C*** and is given by:

$$C(\varepsilon,\gamma,\varepsilon_0) = \frac{2\pi r_e^2 c}{\gamma^2\varepsilon_0} \times \left[2\kappa\ln\kappa + (1+2\kappa)(1-\kappa) + \frac{(4\varepsilon_0\gamma\kappa)^2}{2(1+4\varepsilon_0\gamma\kappa)}(1-\kappa)\right] \tag{E.19}$$

where

$$\kappa = \frac{\varepsilon}{4\varepsilon_0\gamma(\gamma-\varepsilon)} \tag{E.20}$$

The ***allowed kinetic energy range*** for the up-scattered photon energy is

$$\varepsilon \le \varepsilon \le \gamma\frac{4\varepsilon_0\gamma}{(1+4\varepsilon_0\gamma)} \tag{E.21}$$

The ***hard photons can interact with the soft photons*** (and also with soft external thermal photons from the accretion disc) by creating electron-positron pairs, if their CM energy exceeds 1024 eV. This results in an ***attenuation*** of the hard photons and introduces an upper limit for the energy of the hard photons that can escape from the source. This gives a ***hard limit*** on the ***size*** of the emission region (***opacity condition***). For simplicity, we will neglect this effect here.

E.3 Error calculation of the correlation function

The empirical correlation function that has been used to find correlations between the x-ray flux and the γ-flux has been defined as:

$$\begin{aligned}
\rho(\Delta t) &= \frac{\sum_{i,k}\left(F^{\gamma}(t_k = t_i + \Delta t) - \langle F^{\gamma}\rangle\right)\left(F^{x}(t_i) - \langle F^{x}\rangle\right)}{\sqrt{\sum_i \left(F^{\gamma}(t_k = t_i + \Delta t) - \langle F^{\gamma}\rangle\right)^2 \sum_i \left(F^{x}(t_i) - \langle F^{x}\rangle\right)^2}} \qquad \text{(E.22)}\\
&= \frac{\langle xy\rangle - \langle x\rangle\langle y\rangle}{\sigma_x \sigma_y}\\
&\quad with\ the\ shortcuts\ x, y:\\
x_i &= F^{\gamma}(t_k = t_i + \Delta t) \qquad \text{(E.23)}\\
y_i &= F^{x}(t_i) \qquad \text{(E.24)}\\
\delta x_i &= \sigma_{F^{\gamma}(t_k = t_i + \Delta t)} \qquad \text{(E.25)}\\
\delta y_i &= \sigma_{F^{x}(t_i)} \qquad \text{(E.26)}\\
\langle x\rangle &= \frac{1}{N}\sum_i x_i \qquad \text{(E.27)}\\
\langle y\rangle &= \frac{1}{N}\sum y_i \qquad \text{(E.28)}\\
\langle xy\rangle &= \frac{1}{N}\sum_i x_i y_i \qquad \text{(E.29)}\\
\sigma_x^2 &= \frac{1}{N}\sum_i \left(x_i - \langle x\rangle\right)^2 \qquad \text{(E.30)}\\
\sigma_y^2 &\quad \frac{1}{N}\sum_i \left(y_i - \langle y\rangle\right)^2 \qquad \text{(E.31)}
\end{aligned}$$

$\rho = \rho(\Delta t)$ is the empirical correlation coefficient, $F^{\gamma}(t_k)$ is the γ-flux at time t_k, $F^{x}(t_i)$ is the x-ray flux at time t_i, Δt is the time lag between the two datasets and $\sigma_{F^{\gamma}}$ and $\sigma_{F^{x}}$ are the errors on the individual time bin measurement.

The error of the correlation has two components. One part $\sigma_{\rho,flux}$, comes directly from Gaussian error propagation of the errors on the fluxes [WitCom] and the second one $\sigma_{\rho,stat}$ is of statistical nature [Bro90, WitCom] of the correlation. Both of them have to be added. The uncertainty of the correlation coefficient which comes from the errors in the flux measurements are given in terms of the shortcuts above (for space reasons).

$$\begin{aligned}
\sigma_{\rho,tot}^2 &= \sigma_{\rho,flux}^2 + \sigma_{\rho,stat}^2 \qquad \text{(E.32)}\\
\sigma_{\rho,flux}^2 &= \frac{1}{N}\sum_i \left(\frac{\left(y_i - \langle y\rangle\right) - \frac{\sigma_y}{\sigma_x}\left(x_i - \langle x\rangle\right)\rho}{\sigma_x \sigma_y}\right)^2 \delta x_i^2 + \qquad \text{(E.33)}\\
&= \frac{1}{N}\sum_i \left(\frac{\left(x_i - \langle x\rangle\right) - \frac{\sigma_x}{\sigma_y}\left(y_i - \langle y\rangle\right)\rho}{\sigma_x \sigma_y}\right)^2 \delta y_i^2\\
\sigma_{\rho,stat}^2 &= \frac{1-\rho^2}{N-2} \qquad \text{(E.34)}
\end{aligned}$$

N is the number of samples that is being summed up. It is at the order of 80 (for the whole lightcurve in one day bins). The second term $\sigma_{\rho,stat}^2$ is in this case smaller than the first one but has to taken into account as well.

E.4 Re-binning and averaging of flux bins

For several reasons the flux bins must be re-binned or averaged over larger time periods (for correlation calculations or comparisons with other telescopes, for example see Fig. A.41) or flux level bins (for the hardness studies). The averaging has been done by weighting with the observation time of the corresponding run bin.

$$\langle F \rangle = \frac{\sum_{Bin} F_i T_i}{\sum_{Bin} T_i} \tag{E.35}$$

$$\sigma^2_{\langle F \rangle} = \frac{\sum_{Bin} \sigma^2_{F_i} T_i^2}{\left(\sum_{Bin} T_i\right)^2} = \frac{\sum_{Bin} \sigma^2_{F_i} T_i^2}{N_{eq} \sum_{Bin} T_i^2} \tag{E.36}$$

$$N_{eq} = \frac{\left(\sum_{Bin} T_i\right)^2}{\sum_{Bin} T_i^2} \tag{E.37}$$

F is understood as $F = \int (dF/dE)\, dE$. Here $\sigma^2_{\langle F \rangle}$ is variance of the mean and N_{eq} is the so-called equivalent number of events. It can be obtained by error propagation and it is needed to obtain the correct variance of the mean in the case of weighted averages.

Appendix F

Appendix B: Lightcurves and ALPHA plots

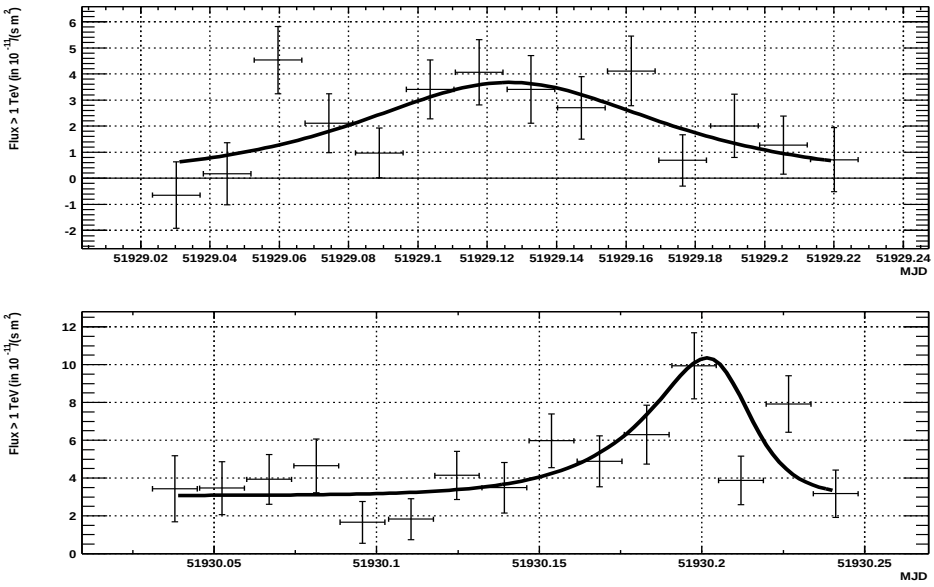


Figure F.1: Lightcurve of the nights 51928, 51929

F.1 The complete daily lightcurve

The following pages present the daily lightcurves of Mkn 421 with more than 7 runs per night from February 2001 until June 2001 as recorded with the CT1 telescope in La Palma, altogether 259 hours of observation. As explained before, a simple flare model

$$F(t) = a + \frac{b}{\left(2^{(t-t_0)/c} + 2^{-(t-t_0)/d}\right)} \tag{F.1}$$

has been fitted to each night, if a simple straight line fit gave a reduced χ^2/NDF worse than 1.5. As starting values have been chosen: a=the constant term from the line fit, b=8.0, c=d=25 minutes, t_0=the highest flux point in the curve. The fastest flares are not covered by the fit. Some nights like 51966 contain flares that are very fast and are significantly outside the flare model.

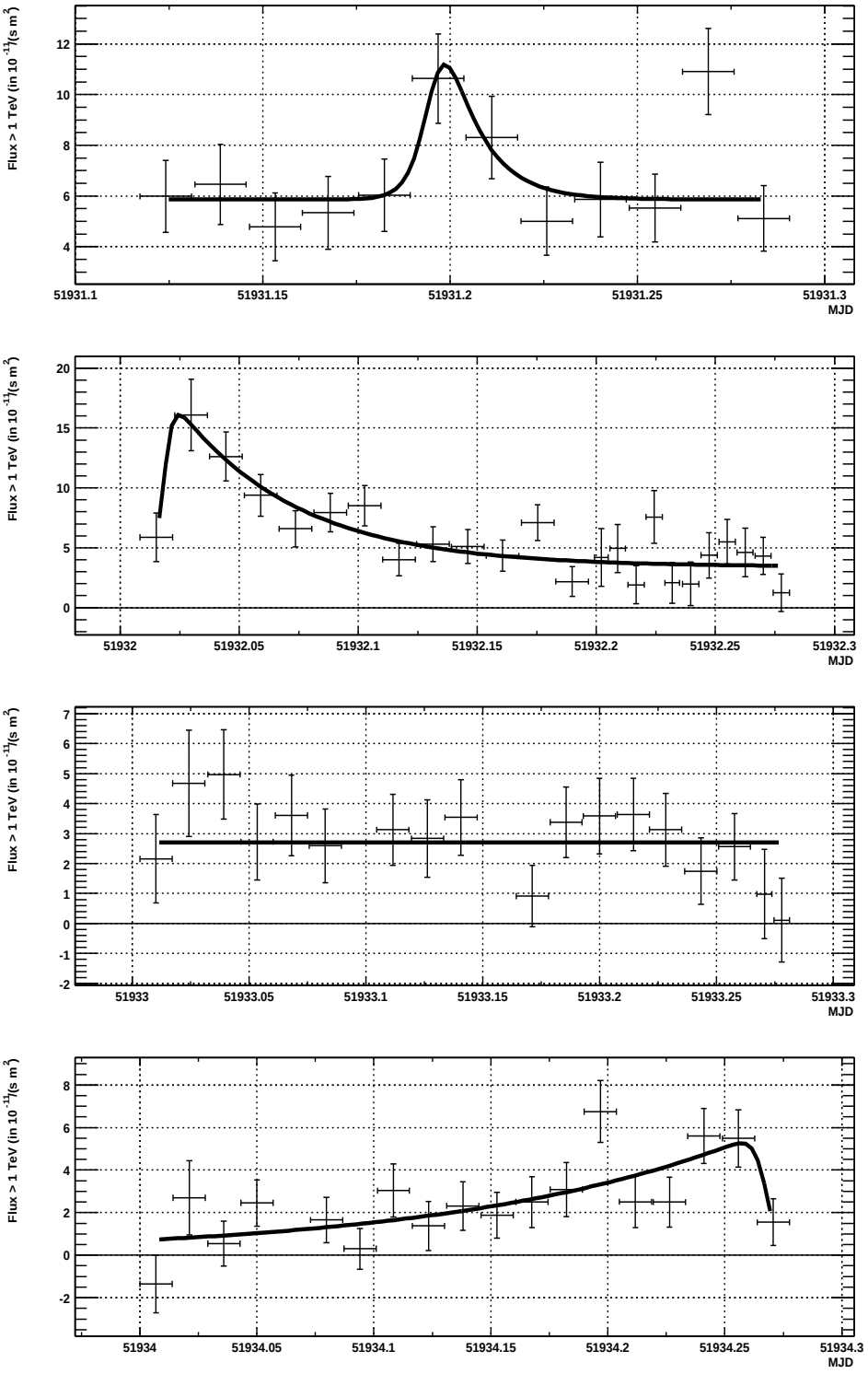


Figure F.2: Lightcurve of the nights 51930, 51931, 51932 and 51933

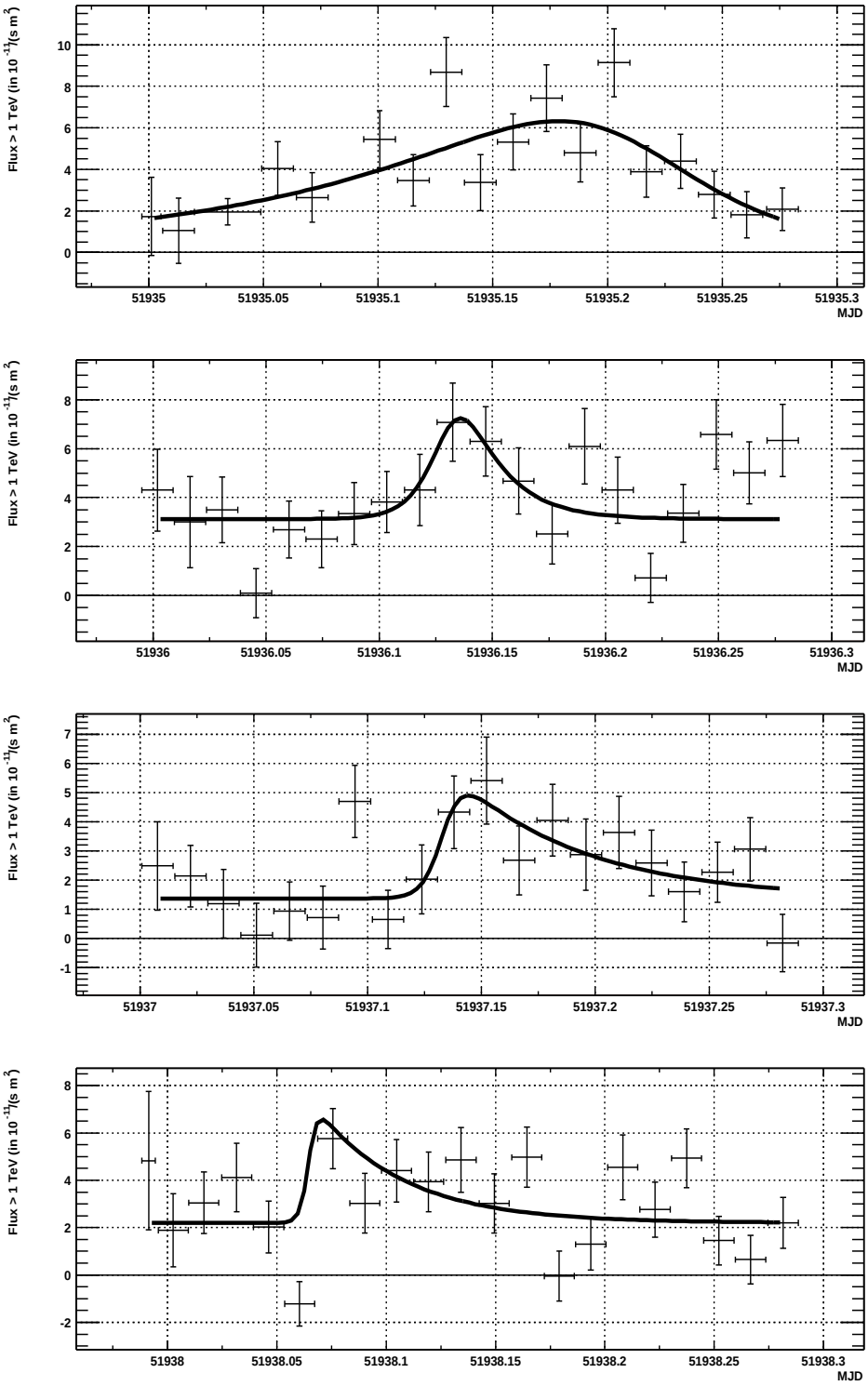


Figure F.3: Lightcurve of the nights 51934, 51935, 51936 and 51937

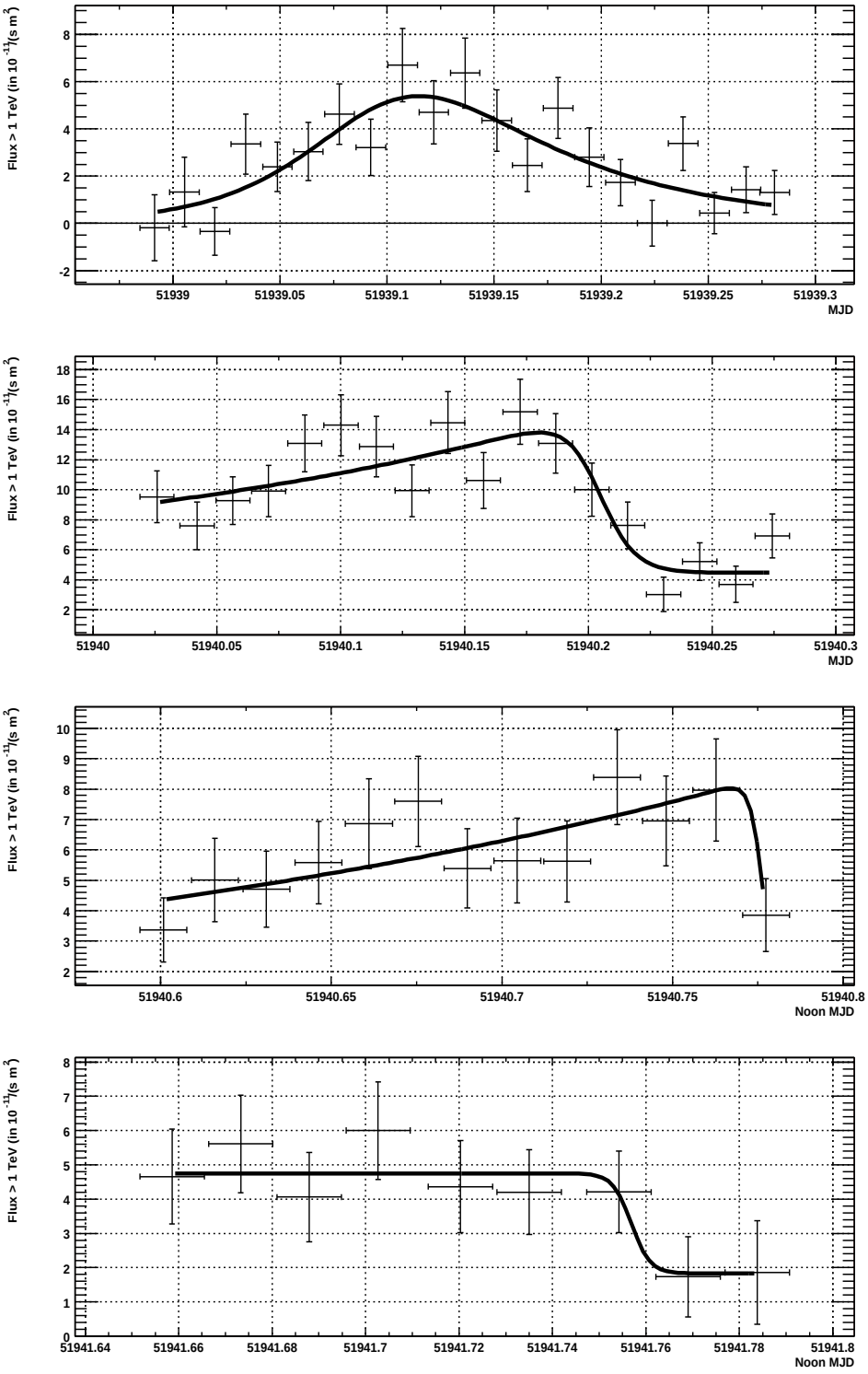


Figure F.4: Lightcurve of the nights 51938, 51939, 51940 and 51941

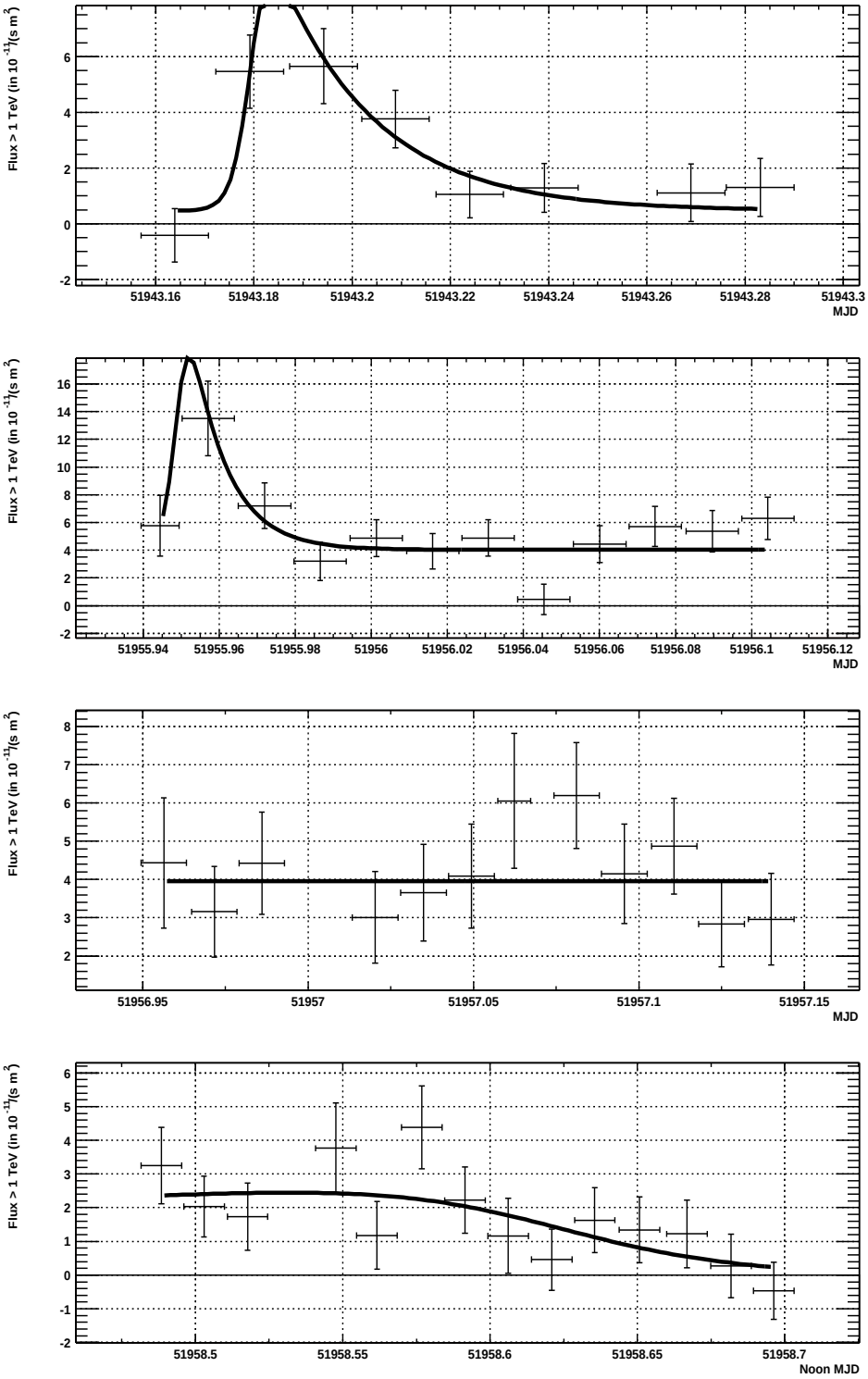


Figure F.5: Lightcurve of the nights 51942, 51928, 51929 and 51930

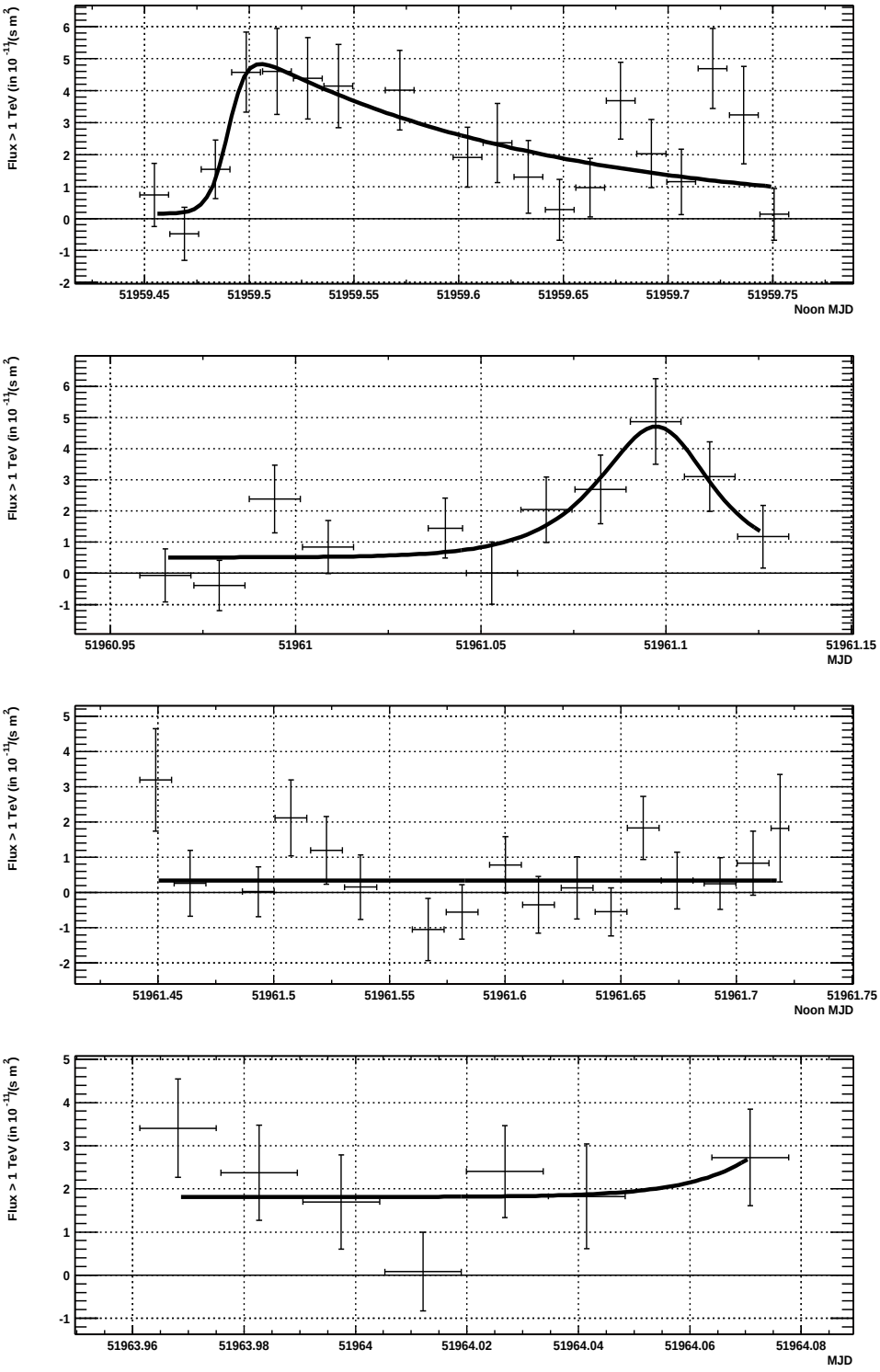


Figure F.6: Lightcurve of the nights 51959, 51960, 51961 and 51963

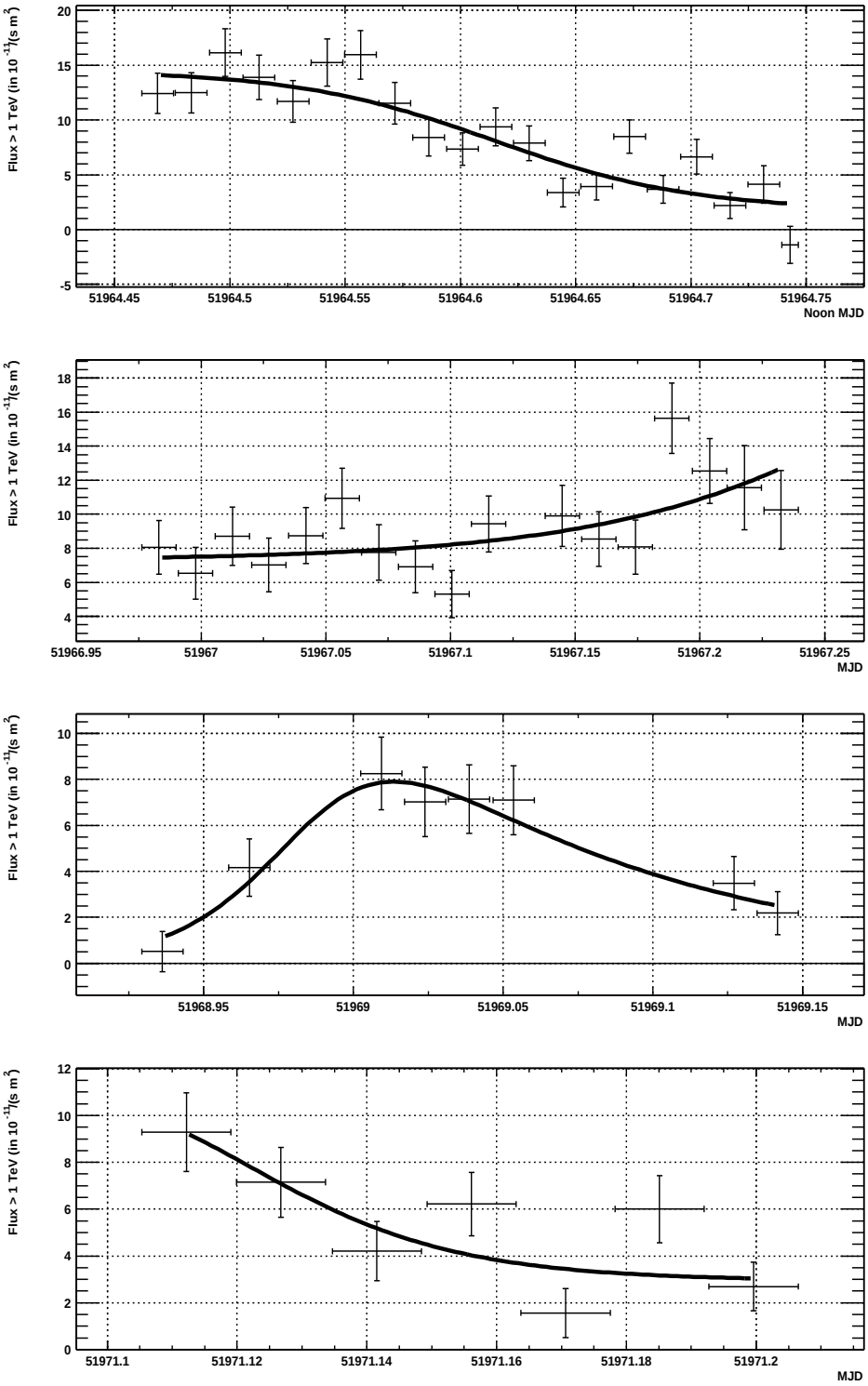


Figure F.7: Lightcurve of the nights 51964, 51966, 51968 and 51970

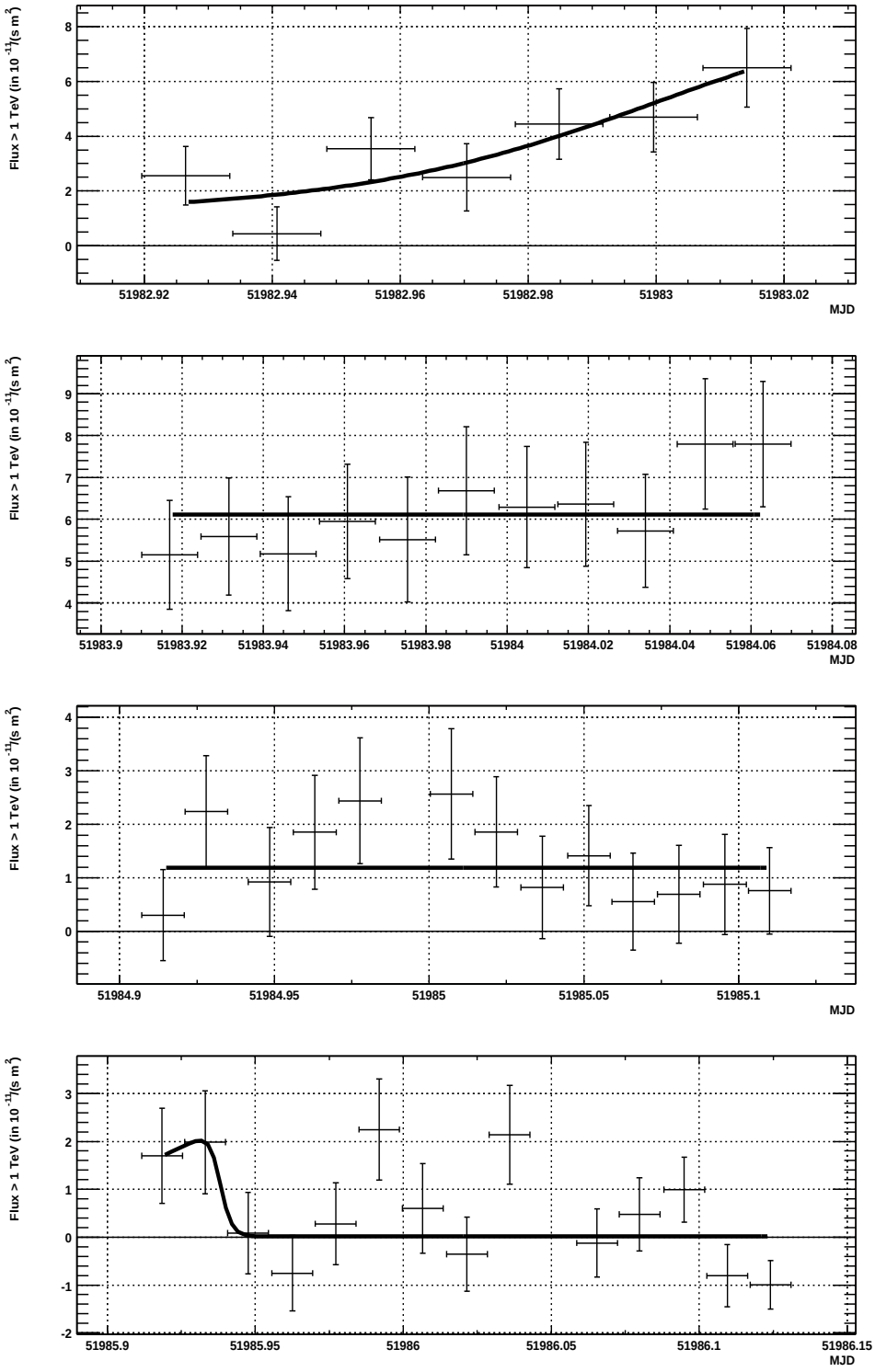


Figure F.8: Lightcurve of the nights 51982, 51983, 51984 and 51985

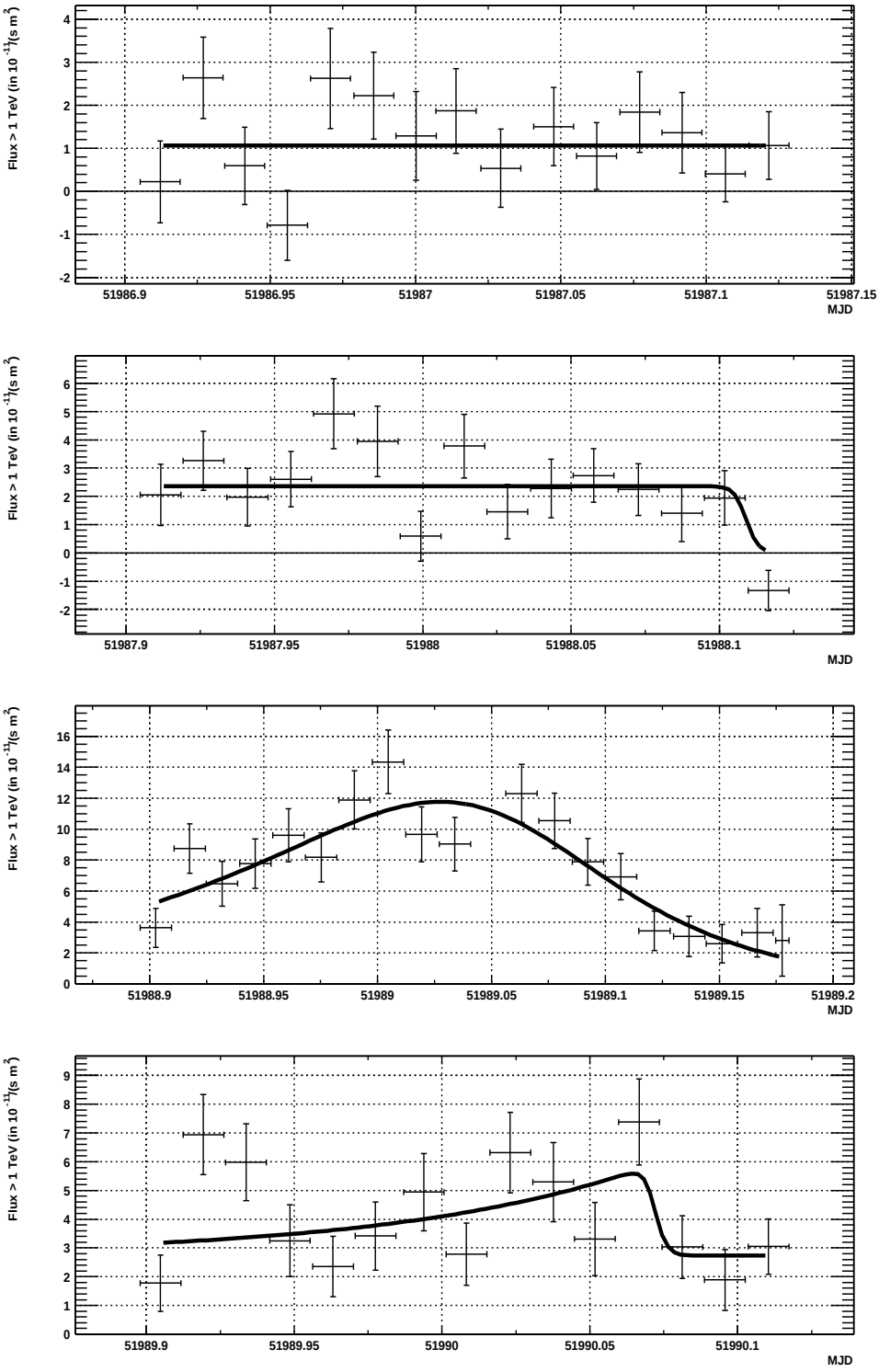


Figure F.9: Lightcurve of the nights 51986, 51987, 51988 and 51989

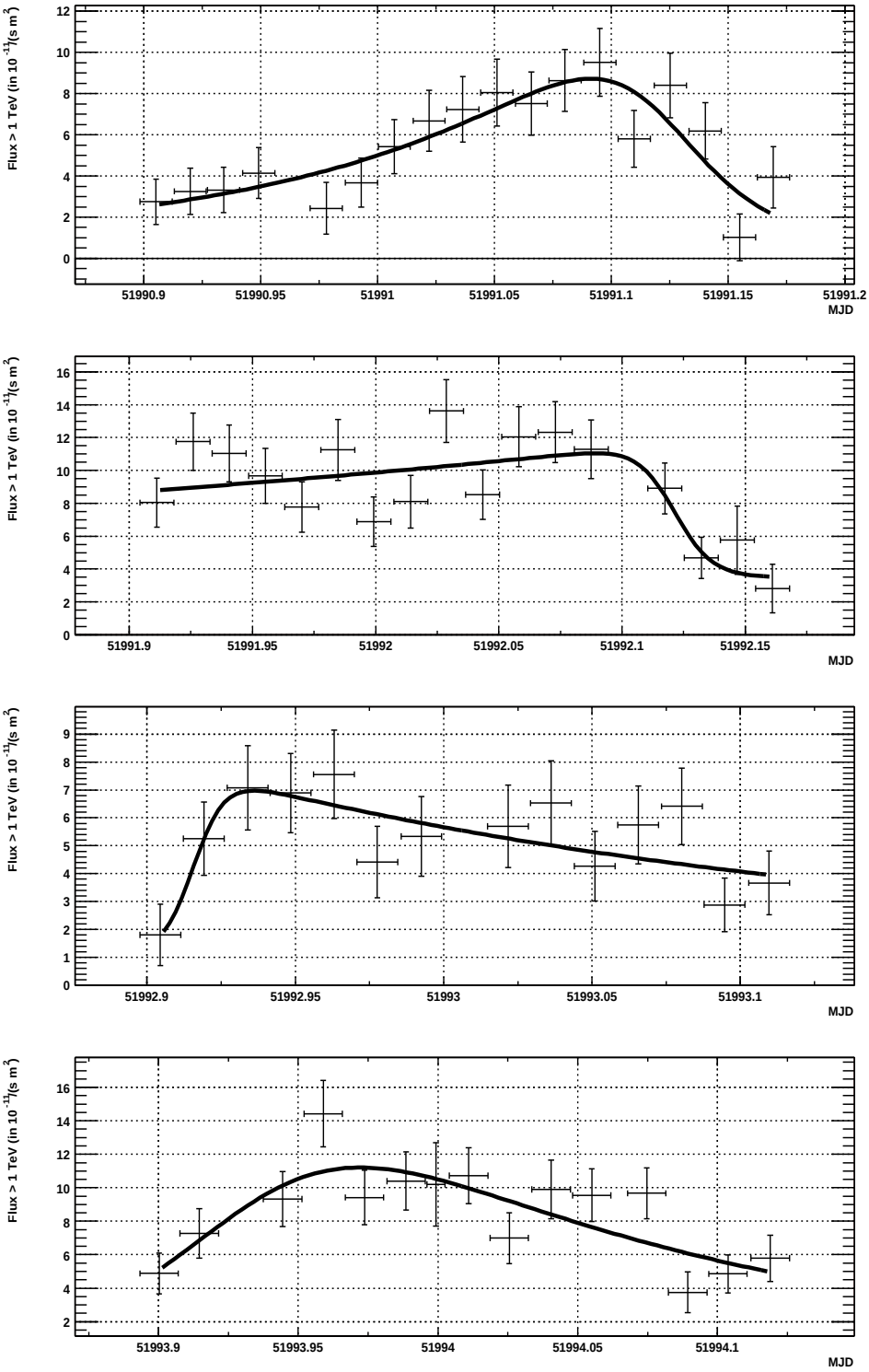


Figure F.10: Lightcurve of the nights 51990, 51991, 51992 and 51993

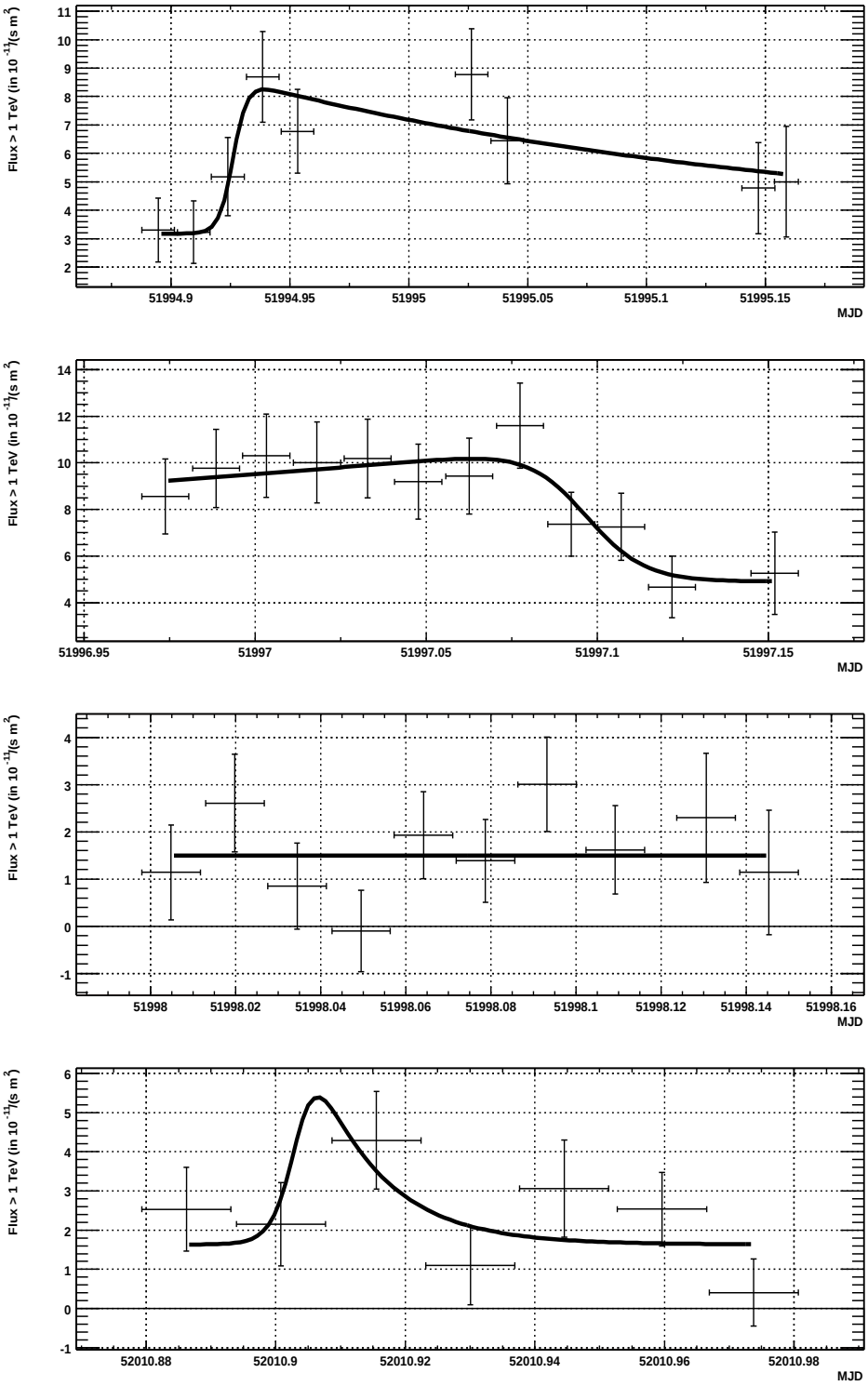


Figure F.11: Lightcurve of the nights 51994, 51996, 51997 and 50210

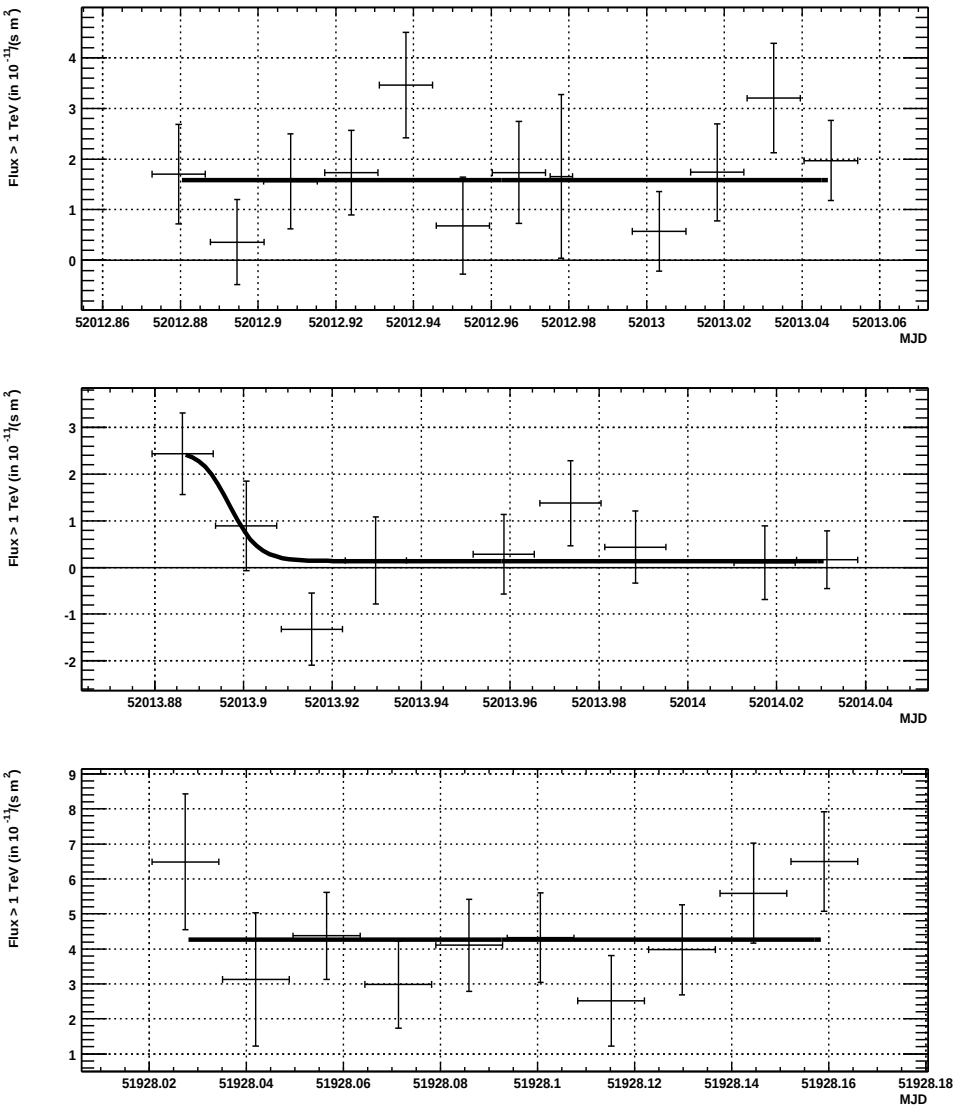
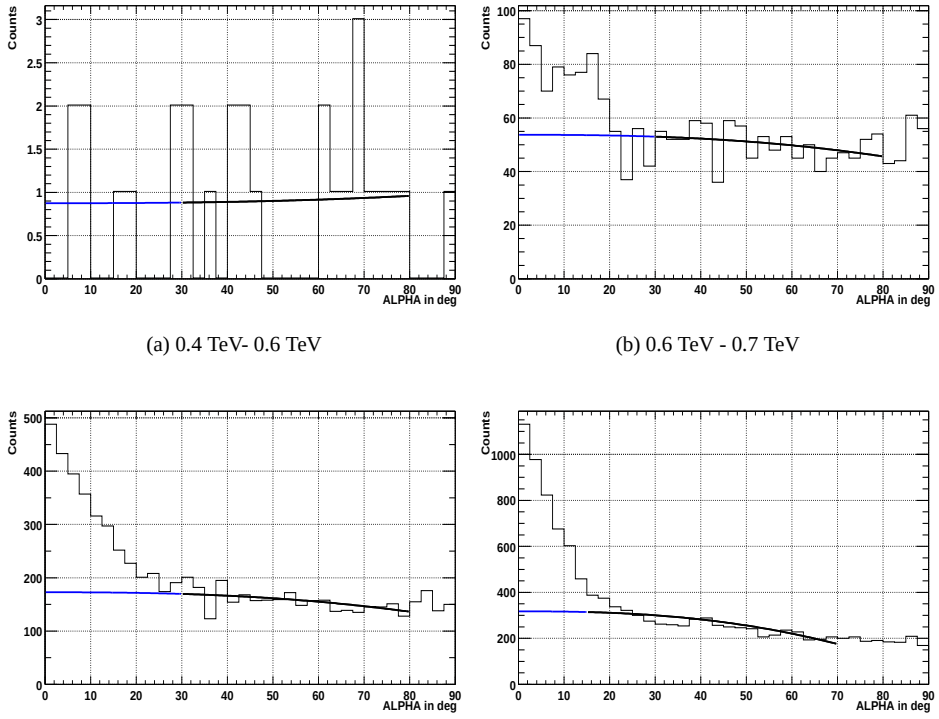


Figure F.12: Lightcurve of the nights 52012, 52013 and 51927



(a) 0.4 TeV- 0.6 TeV

(b) 0.6 TeV - 0.7 TeV

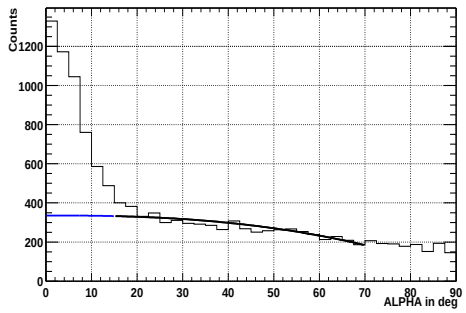
(c) 0.7 TeV - 1.0 TeV

(d) 1.0 TeV - 1.3 TeV

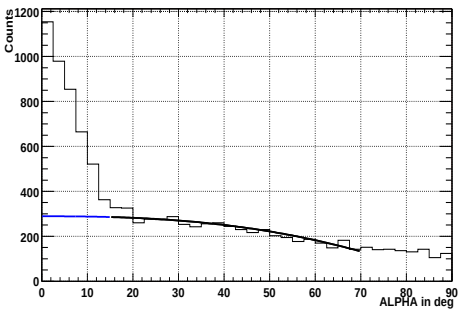
Figure F.13: ALPHA plots for energies from 0.4 TeV to 1.3 TeV

F.2 The estimation of the background for the spectrum

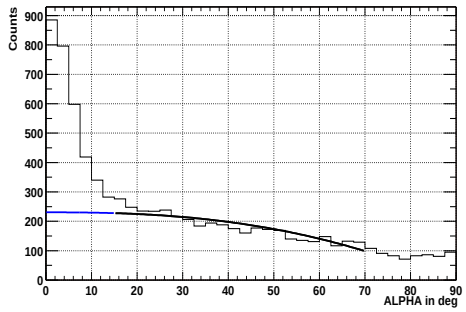
For completeness the plots for the background estimation for the spectrum calculation are shown here. For each energy bin a ALPHA plot has been made and the background from zero up to 18° is estimated by means of a polynomial fit with two free parameters in the ALPHA region without signal. The fit regions have been chosen energy dependent and are: 30°-80° for energies below 1 TeV, 15°-70° for energies from 1 TeV to 5 TeV and 10°-50° for energies above 5 TeV. These values have been chosen to be adequate to the width of the ALPHA distribution of the signal (which becomes wider for lower energies) and to the shape of the background (which becomes more curved for higher energies).



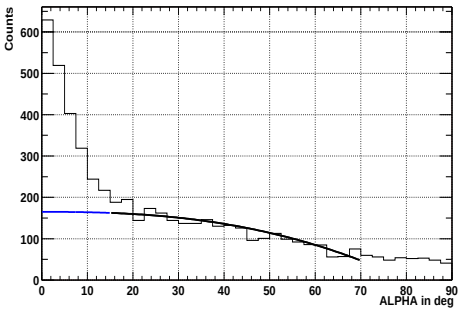
(a) 1.3 TeV - 1.8TeV



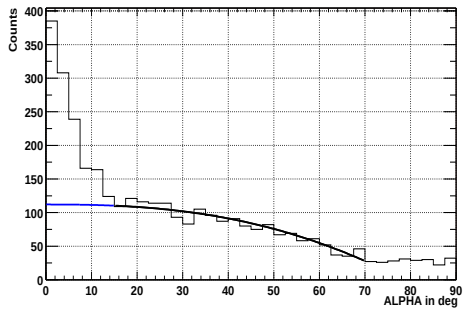
(b) 1.8 TeV - 2.4 TeV



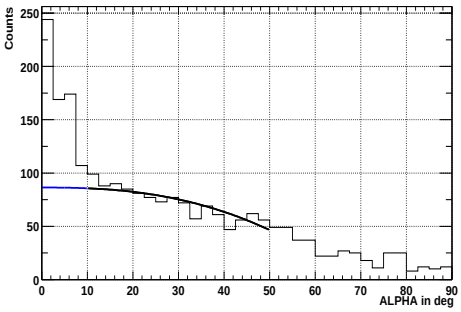
(c) 2.4 TeV - 3.2 TeV



(d) 3.2 TeV - 4.2 TeV

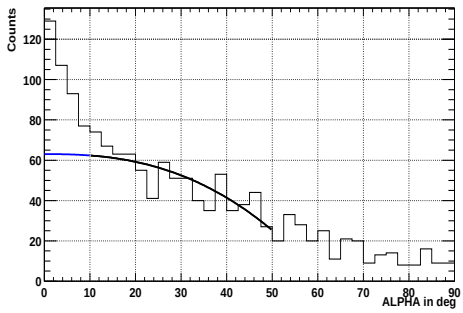


(e) 4.2 TeV- 5.6 TeV

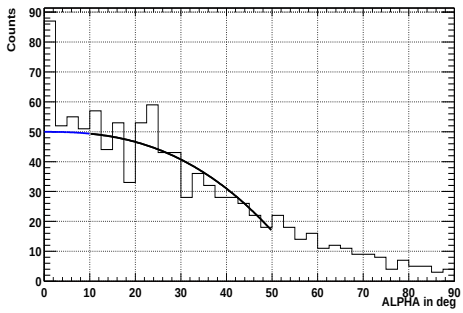


(f) 5.6 TeV - 7.5 TeV

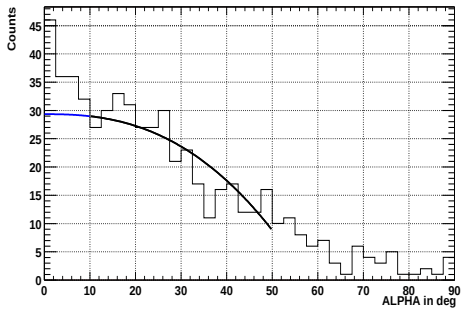
Figure F.14: ALPHA plots for energies from 1.3 TeV to 7.5 TeV



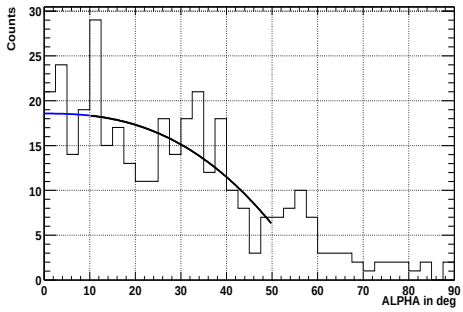
(a) 7.5 TeV - 10.0 TeV



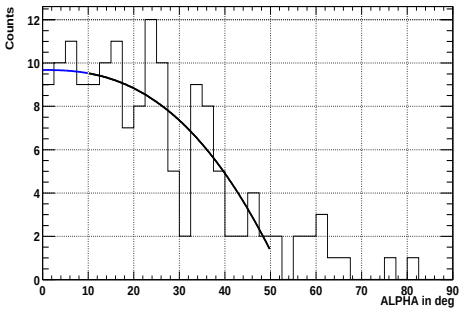
(b) 10.0 TeV - 13.3 TeV



(c) 13.3 TeV- 17.8 TeV



(d) 17.8 TeV - 23.7 TeV



(e) 23.7 TeV - 31.6 TeV

Figure F.15: ALPHA plots for energies from 7.5 TeV to 31.6 TeV

Bibliography

[Ame00] Giovanni Amelino-Camelia and Tsvi Piran, Cosmic rays and TeV photons as probes of quantum properties od space time, hep-ph/0006210, June 2000

[Ame96] G. Amelio-Camelia et al., Distance Measurement and Wave Dispersion in a Liouville-String Approach to Quantum Gravity, May 1996

[Aha00] Aharonian et al., Complex Spectral Variability from Intensive Multi-wavelength Monitoring of Mkn 421 in 1998, astroph/0008505, 2000

[Aha99/1] Aharonian, Astrophysics and Astronomy, 349, 29, 1999

[Aha99/2] Aharonian, Astrophysics and Astronomy, 350, 757, 1999

[Bad98] Badley, Peterson et al., On uncertainties in Cross-Correlation Lags and the Reality of Wavelength dependent Continuum Lags in Active Galactic Nuclei, Astronomical Society of the Pacific, 110, pp. 660-670, June 1998

[Bir02] J.A. Biretta, W. Junor, M. Livio, Evidence for initial jet formation by an accretion disc in the radio galaxy M87, New atronomy reviews 46

[Bed96] W. Bednarek, R. Protheroe, Modulation of AGN gamma rays by interaction with X-rays from an accretion disc hot spot, astro-ph/9612211, Dec 96

[Beg80] Begelman et al., Nature 287, 307, 1980

[Bro90] Bronstein, Handbuch der Mathematic, 1990

[Bla01] O. Blanch, M. Martinez, Exploring the gamma ray horizon with the next generation of gamma ray telescopes, astro-ph/0107582, July 2001

[Blo96] S. D. Bloom and A. P. Marscher, An analysis of the Synchrotron Self-Compton Model for the Multi-Wave band spectra of blazars, Astrophysical Journal, 461: 657+, April 1996

[Böt97] M. Böttcher, Zeitabhängiger Strahlungstransport in Jets von aktiven galaktischen Kernen, Ph.D. Thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, January 97

[Bra93] H. Bradt, X-ray timing explorer mission, Astronomy and Astrophysics, Supplement Series 97, 355-360, January 93

[Car99] M.J. Carson, B. McKernan, T.Yaqoob, D.J. Fegan, Short time scale variability in the broadband emission of the blazars Mkn 421 and Mkn501, astro-ph/9906200, June1999

[CAT89] S. Bohec, A new analysis method for very high definition Imaging Cherenkov Telescopes as applied to the CAT telescope, astro-ph/9804133, April 98

[CAT01] Temporal and spectral gamma-ray properties of Mkn 421 above 250 GeV from CAT observations between 1996 and 2000, astro-ph/0106196, June 2001

[Car99] M. Carson, et al., SHort-timescale Variability in the broadband emission of the Blazars Mkn 421 and Mkn501, astro-ph/9906200, January 1999

[Dar97] A. Dar and A. Laor, Hadronic production of TeV Gamma-ray flares from Blazars, The astrophysical journal, 478:L5-L8, March 1997

[Des95] G. Zech, Comparing statistical Data to Monte Carlo Simulation-Parameter, Fitting and Unfolding, Desy 95, 13, June 1995

[Fab97] B. Fabbro, Linear Discriminant analysis with stepwise method. Application for MSSM Higgs boson search in the $hA \rightarrow b\overline{b}b\overline{b}$ channel at LEP2, ALEPH 97-012, February 1997

[Fer49] E. Fermi, On the origin of the Cosmic radiation, Physical Review, vol 75, number 8, April 1949

[Fis37] R. Fisher, The use of multiple measurement in taxonomic problems. Annals of Eugenics, Vol.7, 1936

[Gai90] Thomas K. Gaisser, Cosmic Rays and Particle Physics, Cambridge University press 1990

[Gal02] Y. A. Gallant, Particle acceleration at relativistic shocks, astro-ph/0201243, January 2002

[Gru96] D. Gruber, The high energy X-ray timing experiment on XTE, Astronomy and Astrophysics, Supplement Series 120, C641+, December 96

[Gui98] G. Guisellini, Blazars: Recent developments, astro-ph/9810230, October 98

[Hau01] M. G. Hauser, The cosmic infrared background: Measurements and implications, astro-ph/0105539, May 01

[Hil96] A. Hillas, Differences between Gamma-Ray and Hadronic Showers, Space Science Review, 75: 17-30, January 1996

[Hil85] A. Hillas, Cherenkov light images of EAS produced by primary gamma, In NASA, Goddard Space Flight Center 19th Intern. Cosmic Ray Conference, Vol 3, 445-448, August 85

[Hil98] Hillas et al., TeV Blazars: Status of Observations, astroph/9812029, 1998,

[Hil82] A. Hillas, Angular and Energy Distributions of charged particles in electron-photon cascades in air, J. Phys. G. 8, 1461-1473, 1982

[Hsu02] S. Hsu, A Laboratory Plasma Experiment for Studying Magnetic Dynamics of Accretion Discs and Jets, astro-ph/0202380, February 02

[Ino96] S. Inoue and F. Takahara, Electron acceleration and gamma-ray emission from blazars, Astrophysical journal, 463: 555-564, June 1996

[Jor01] M. Jordan, J. H. Buckley, R. Cowsik et al., Multiwavelength observations of Mkn 421, Proceedings of ICRC 2001, 2001

[Jel58] J. V. Jelley. Cherenkov radiation and its applications. Pergamon Press,1958

[Kat01] J. Kataoka, et al., Five year monitoring of Tev Blazars with ASCA and RXTE, astro-ph/0105029, May 2001

[Kes01] M. Kestel, A method to correct HILLAS Parameters of Imaging Cherenkov Telescope Data taken at different Background Light levels, Proc. 27th ICRC (Hamburg), OG269, 2001

[Kne01] T. Kneiske, K. Mannheim and D. Hartmann, Evolving Stellar Background Radiation and Gamma-Ray Optical Depth, astro-ph/0011013

[Kne02] T. Kneiske, Implications of cosmological gamma-ray absorption, I. Evolution of the metagalcic radiation field, Astronomy & Astrophysics 386, 1-11, 2002

[Koi02] S. Koide et al., 3-D General relativistic MHD simulatins of Generation Jets, astro-ph/0202396, February 2002

[Kra01] D. Kranich, Temporal and spectral characteristics of the active galactic nucleus Mkn 501 during a phase of high activity in the TeV range, PhD thesis at the Technische Universität München, March 2001

[Kre02] F. Krennrich, Discovery of Spectral Variability of Mkn 421 at TeV Energies, astro-ph/0207184, July 2002

[Kud99] T. Kudoh, S. Aoki, S. Koide and K. Shibata, Are blazar jets magnetically driven outflows ?, Astronom. Nach. 320, 4/5, 1999

[Mar99] Alan Marscher, The compact jets of TeV blazars, Astroparticla physics 11, 19-25, 1999

[Lev96] A. Levine, H.Bradt, W. Cui et al., First results from the All-Sky Monitor on the Rossi X-Ray Timing Explorer, Astrophysical Journal Letters, 469: L33+, September 96

[Liv02] Mario Livio, The jet set, Nature, vol. 417, May 2002

[LiMa83] Ti-Pei Li and Yu-Qian Ma, Analysis Methods for Results in Gamma-Ray Astronomy, The astrophysical Journal, 272:317-324, September 83

[LON97/1] Longair, High energy astrophysics, volume 2, p340, 1997

[LON97/2] Longair, High energy astrophysics, volume 2, p326, 1997

[Lon1/92] M. Longair, High energy astrophysics, vol. 1, p. 122-p. 127, 1992

[Lon95] J. Lonway, J. Wobel, A helical jet in the orthogonally misaligned BL LAC object Mkn 501, Astrophysical journal, 439, 98-112, January 1995

[Lor01] E. Lorenz, Status of ground-based GeV/TeV gamma ray astronomy, Journal of physics G:Nuclear and particle physics, 27, 1675-1690, 2001

[MagCom] Norbert Magnussen, private communication

[Mag01] Norbert Magnussen, Course given at IFAE, July 2001(2002) 239-245, 2002

[Mar02] Alan Marscher et al., Observational evidence for the accretion-disk origin for a radio jet in an active galaxy, Nature vol 417, 6 June 2002

[Man00] F. M. and K. Mannheim. A possible black hole binary in Mkn 501, Proc. of the Heidelberg International Symposium on High Energy Gamma-Ray Astronom, 2000

[Mas01] Masterson et al., Temproral ans spectral gamma-ray properties of Mkn 421 above 250 GeV from CAT observations between 1996 and 2000, astro-ph/010696, 2001

[Mas99] A. Mastichiadis, J. Kirk, Models of Variability in Blazar Jets, astro-ph/9903280, March 99

[Mei01] D. Meier, Magnetohydrodynamic Production of Highly Relativistic Jets, Proceedings of Astrophysical Phenomena Revealed by space VLBI, 2001

[Mir94] R. Mirzoyan, R. Kankanian, P. Sawallisch et al., The first Imaging Air Cherenkov telescope of the HEGRA collaboration, Nuclear instruments and Methods, A 351:513+, 1994

[Mir00] Mirzoyan, Conversion factor calibration for MAGIC based on the use on measured F-factor of PMTs, MAGIC internal note, April 2000

[MirCom] Razmik Mirzoyan, MPI für Physik, Munich, private communication

[Nis02] K. Nishikawa, 3-D General Relativistic MHD Simulations of Generation jets, astro-ph/0202396, February 2002

[Ose01] S. Oser, et al., High energy Gamma-ray Observations of the Crab Nebula and Pulser with the Solar Tower Atmospheric Cherenkov Effect Experiment, Astrophysical Journal, 2001

[Pad95] P. Padovani, Unified Schemes for Radio-Load Active Galactic Nuclei, The Publications of the Astronomical Society of the Pacific 107: 803+, September 1995

[Pad1/99] P. Padovani, Unified Schemes and the two Classes of BL LACs, astro-ph/9901128, Jan 99

[Pad2/99] P. Padovani, High Energy Emission from AGN and Unified Schemes, astro-ph/9901130, Jan99

[Pet97] D. Petry, Beobachtung hochenergetischer γ-Strahlung (E>1 TeV) aus Richtung der aktiven Galaxien Mkn 421 und Mkn 501, PhD. thesis, Technische Universität München, Föhringer Ring 6, 80805 München, September 1997, MPI-PhE/97-27

[Pir01] F. Piron et al., Temporal and spectral gamma-ray properties of Mkn 421 above 250 GeV from CAT observations between 1996 and 2000, astro-ph/0106196, June 2001

[Qui96] Quinn J et al, Astrophys. J. 456, L83, 1996

[Rau95] G. Rauterberg et al., A new 127 Pixel Camera for Cherenkov Telescopes of the HEGRA Array on la Palma, In 24th Interbational Cosmic Ray Conference (Rome), volume 3, p 460+, 1995

[Rey93] P. T. Reynolds, C. W. Akerlof, M. F. Cawley et al., Survey of candidate gamma-rays sources at TeV energies using a high resolution Cherenkov imaging system - 1988-1991, Astrophysical Journal, 404: 206-218, February 1993

[Sal98] M. Salvati, Rapid Variability of Gamma-Ray Blazars: A model for Markarian 421, astro-ph/9801049, Jan 98

[Sch01] Schweizer, The calibration of the MAGIC telescope camera, to be published in IEEE Trans. Nucl. Sci., 2001

[Sal98] M. Salvati et al., Rapid variability of gamma-ray blazars: a model for Mkn 421, astro-ph/9801049, January 1998

[Sch00] T. Schweizer and M. Dosil, Measurements of the NSB in La Palma on the Roque de los muchachos, September 2000

[Ste95] F. W. Stecker, O. C. De Jager, Absorption of high energy gamma rays by low energy intergalactic photons, astro-ph/9501065, Jan 95

[Ste01] F. W. Stecker, O. C. De Jager, Extragalactic gamma ray Absorption and the intrinsic spectrum of Mkn 501 during the 1997 flare, astro-ph/0107103, October 01

[Tan98] Tanimori et al., TeV gamma-ray observations of southern BL LACs with the CANGAROO 3.8m imaging Telescope, astro-ph/9811260, 1998

[Tav98] F. Tavecchio and Lausra Marashi, Constarints on the Physical Parameters of Tev Blazars, astrop-ph/9809051, 1998

[Tak96] T. Takahashi, et al., Ap J, 470, L89, 1996

[TW] This thesis

[Vas99] V. V. Vassiliev, Extragalactic background light absorption signal in the TeV γ-ray spectra of Blazars, astro-ph/9908088, August 99

[Wan02] Y. Wang ad P. Biermann, Efects of galaxy mergers on the faint source counts and the backgound, to be published in Astronomy and astrophysics

[Wie94] B. Wiebel, Chemical Composition in High Energy Cosmic Rays, Technical Report, Bergische Universität Gesamthochschule Wuppertal, January 1994

[WitCom] Wolfgang Wittek, MPI für Physik, Munich, private communication

[Zwe97] J. A. Zwerink, et al., The TeV Gamma-ray sectrum of Markarian 421 during an intense flare, The astrophysical journal, 490, L141-L14, December 1997