Measurement of the energy spectrum of the BL Lac object PG1553+113 with the MAGIC Telescope in 2005 and 2006

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von

Herr Dipl.-Phys. Thomas Hengstebeck
geboren am 19.08.1974 in Siegen

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Christian Limberg

Gutachter:

1. Prof. Dr. Thomas Lohse
2. Prof. Dr. Claus Grupen
3. Prof. Dr. Hermann Kolanoski

eingereicht am: 4. Oktober 2006
Tag der mündlichen Prüfung: 1. März 2007

Abstract

In this thesis new data analysis methods for the MAGIC experiment were implemented, which are especially suited for the investigation of low energy γ-ray events. They were successfully tested by means of Monte Carlo studies and applied to observational data of the Crab Nebula and of the extragalactic γ-ray source PG1553+113.

These methods extend from image cleaning techniques and the utilization of new image parameters to sophisticated g/h-separation and energy estimation approaches. For the first time in γ-ray astrophysics the advantages of classification and regression trees were exploited in order to improve existing 'classical' methods.

The work of this thesis can be grouped into a methodological, programming-technical and an analysis part:

Methodological and programming-technical work

- Setup of a stand-alone process-oriented analysis framework for the MAGIC data analysis. Automation of the analysis steps from the download of the raw data up to the calculation of the image parameters.

- Participation in the development of a new image cleaning method (see chapter 4) optimized in order to achieve an efficient signal extraction and night sky background (NSB) noise suppression. Here, particularly the implementation of an incremental next-neighbor pixel search logic shall be noted.

- Calculation of the FADC pedestal and the pedestal RMS directly from the data (no dedicated pedestal runs needed, see chapter 4). In contrast to the standard Mars software [BWM03] the pedestals and the RMS are applied to exactly the data stream they were extracted from, which requires two data processing loops.

- Implementation of a continuous calibration correction which runs parallel to the data processing.

- Development and implementation of new image parameters:

  Exploitation of the asymmetries of the Cherenkov light image's charge and time distributions in order to derive new image parameters, which amongst others provide a head-tail information of the shower.

  Programming of a recursive algorithm for the island detection (separated areas in the Cherenkov light image).

  Testing of these image parameters in Monte Carlo studies.

- Utilization of the Random Forest classification and regression methods for the g/h-separation and the energy estimation (see chapter 5). Significant improvement of the g/h-separation in comparison with a 'Scaled Hillas Parameters' standard technique. The simplicity and good performance mainly of the RF g/h-separation (see section 5.4) led to an implementation and intensive usage of the RF method also in the standard analysis of data taken with the MAGIC telescope.

- Modification of an algorithm for the calculation of the very high energy (VHE, see section 1.1) γ-ray absorption at the extragalactic background light provided by [KNS03] which follows the approach of [Vas99b]. The modifications include the implementation of an up to date model of the EBL background energy density distribution and the extension of the algorithm validity to redshifts $z > 0.3$.

- Completion of the SSC model code provided by [K$^+$03] by a minimization procedure based on the TMinuit package (see [BR06]). Implementation of the photon-photon absorption algorithm mentioned above in order to obtain an inverse Compton γ-ray spectrum, which can be directly compared to the experimental data.

Analysis work

- Adjustment of the Monte Carlo parameters point spread function (PSF) and mirror reflectivity.

- Measurement of a differential energy spectrum of the Crab Nebula from observational data recorded by the MAGIC telescope in September 2004. Confirmation of the flux level and the spectral index in the energy range $300\,\mathrm{GeV} < E < 1\,\mathrm{TeV}$ as measured by the HEGRA experiment [A$^+$04], whereby the 'HEGRA measured flux' (range $500\,\mathrm{GeV} \lesssim E \lesssim 86\,\mathrm{TeV}$) was extrapolated down to $300\,\mathrm{GeV}$. Confirmation of the differential energy spectrum of the Crab Nebula as measured by [WM05] in an independent MAGIC analysis for energies $100\,\mathrm{GeV} < E < 1\,\mathrm{TeV}$.

  Stable results of the presented analysis for widely varying g/h-separation cuts could be achieved.

- Measurement of the γ-ray emission of the BL Lac PG1553+113 by means of observations carried out by the MAGIC telescope in 2005 and 2006 with high significance ($> 8\sigma$).

- Determination of a differential energy spectrum of PG1553+113 in the energy range $80\,\mathrm{GeV} < E < 600\,\mathrm{GeV}$. General agreement with results from an independent analysis [A$^+$06c] and H.E.S.S. measurements [A$^+$06b].

The analysis procedure - tested on Monte Carlo data - was demonstrated to be reliable in the investigation of the Crab Nebula γ-ray emission yielding a significant excess in the energy range below 100 GeV in only 1.7 h observation time.

The analysis of data taken on the BL Lac PG1553+113 yielded significant excesses for both years 2005 and 2006. The combined *alpha* histogram shows a signal in excess of 8σ. In the further analysis a spectrum could be derived for the combined data sets of 2005 and 2006 (integral flux above 200 GeV: $F_{>200\,\text{GeV}} = (1.7 \pm 0.3_{stat}) \cdot \frac{10^{-12}}{\text{cm}^2\,\text{s}}$, power-law index: $\Gamma = 3.6 \pm 0.3_{stat}$). This spectrum was used to constrain the redshift z of PG1553+113 with the result $z < 0.68$ (2σ confidence level). A simple SSC (synchrotron self-Compton) model is able to fit the broad band spectral energy distribution (SED) extending from the optical to the γ-ray region reasonably well, especially if a low EBL density level is assumed. As expected from the multi-wavelength investigations of various blazars, a ratio η between e^--energy density and magnetic field energy density significantly larger than 1 was obtained.

Zusammenfassung

In dieser Doktorarbeit wurden im Rahmen des MAGIC Experimentes neue Datenanalysemethoden implementiert, die sich insbesondere für die Analyse von Ereignissen niedriger Gammastrahlungsenergie eignen. Die Methoden konnten erfolgreich in Monte Carlo Studien getestet und auf Beobachtungsdaten des Krebsnebels und der extragalaktischen Gammastrahlungsquelle PG1553+113 angewandt werden.

Diese Methoden reichen von 'image cleaning' Techniken und der Nutzung neuer Bildparameter bis zu fortgeschrittenen g/h-Separations- und Energieabschätzungsverfahren. Zum ersten Mal wurden die Vorteile von Klassifikations- und Regressionsbäumen in der Gamma-Astrophysik ausgenutzt, um existierende 'klassische' Methoden zu verbessern.

Die Arbeitsgebiete dieser Dissertation lassen sich in einen programmiertechnisch methodischen und einen analytischen Bereich einteilen:

Methodische und programmiertechnische Arbeit

- Entwicklung einer unabhängigen prozessorientierten Programmumgebung zur MAGIC Datenanalyse. Automatisierung der Analyseschritte vom 'Download' der Rohdaten bis zur Berechnung der Bildparameter.

- Beteiligung an der Entwicklung einer neuen 'image cleaning' Methode (siehe Kapitel 4), die mit der Zielsetzung optimiert wurde, eine effiziente Signalextraktion mit Unterdrückung des 'night sky background' (NSB) Rauschens zu erreichen. Hier sei insbesondere die Implementierung einer inkrementalen 'next neighbor'-Pixel-Suchlogik genannt.

- Berechnung des FADC Grundniveaus und der zugehörigen Varianz ('pedestal' und 'pedestal RMS') direkt von den zu bearbeitenden Daten (keine speziellen Datennahme-'Runs' nötig). Im Unterschied zur Standardsoftware Mars [BWM03] werden 'pedestal' und 'pedestal RMS' auf genau den Datenbereich angewandt, der auch zu ihrer Berechnung verwandt wurde, was zwei Schleifen in der Datenverarbeitung erfordert.

- Implementierung einer kontinuierlichen Kalibrierungskorrektur, die während der Datenverarbeitung läuft.

- Entwicklung und Implementierung neuer Bildparameter:

  Ausnutzung der Asymmetrien von Ladungs- und Zeitverteilung des Cherenkovlicht-bildes zur Berechnung neuer Bildparameter, die unter anderem eine 'head-tail' In-formation des Schauerbildes liefern.

  Programmierung eines rekursiven Algorithmus zur Bestimmung von 'islands' (sepa-rierte Bereiche im Cherenkovlichtbild).

  Test dieser Parameter in Monte Carlo Studien.

- Nutzung der 'Random Forest' Klassifikations- und Regressionsmethoden in der g/h-Separation und der Energieabschätzung (siehe Kapitel 5). Signifikante Verbesserung der g/h-Separation im Vergleich zu einer 'Scaled Hillas Parameter' Standardtech-nik. Die Einfachheit und gute Performance hauptsächlich der RF g/h-Separation (siehe Abschnitt 5.4) führte zur Implementierung und intensiven Anwendung der RF Methode auch in der Standard MAGIC Datenanalyse.

- Modfizierung eines Algorithmus zur Berechnung der 'very high energy' (VHE, siehe Abschnitt 1.1) Gammastrahlungsabsorption durch extragalaktisches Hintergrund-licht (EBL), der von [KNS03] zur Verfügung gestellt wurde und dem Ansatz von [Vas99b] folgt. Die Modifikationen beinhalten die Implementierung eines aktuellen EBL-Modells sowie die Erweiterung der Gültigkeit des Algorithmus auf Rotverschie-bungen $z > 0.3$.

- Komplettierung eines SSC-Modell-Quellcodes [K$^+$03] durch eine auf dem TMinuit Paket (siehe [BR06]) basierenden Minimierungsprozedur. Implementierung des oben erwähnten Algorithmus zur Berechnung der Photon-Photon Absorption, um ein 'in-verse Compton' Gamma-Spektrum zu erhalten, das direkt mit den experimentellen Daten verglichen werden kann.

Analysearbeit

- Abstimmung der Monte Carlo Parameter PSF und Spiegelreflektivität.

- Bestimmung des differentiellen Energiespektrums des Krebsnebels aus Beobach-tungsdaten, die mit dem MAGIC Teleskop im September 2004 aufgenommen wur-den. Bestätigung von Flusslevel und spektralem Index im Energiebereich $300\,\mathrm{GeV} < E < 1\,\mathrm{TeV}$, wie zuvor mit dem HEGRA Experiment bestimmt, wobei der vom HE-GRA Experiment gemessene Fluss (Bereich $500\,\mathrm{GeV} \lesssim E \lesssim 86\,\mathrm{TeV}$) bis nach $300\,\mathrm{GeV}$ extrapoliert wurde. Bestätigung des differentiellen Energiespektrums des Krebsnebels wie bereits von [WM05] in einer unabhängigen MAGIC Analyse ge-messen.

  Es wurden stabile Resultate der präsentierten Analyse für sehr verschiedene g/h-Separationsschnitte erreicht.

- Messung der Gamma Emission des BL Lac Objekts PG1553+113 mittels Beobachtungen des MAGIC Teleskops in den Jahren 2005 und 2006 mit hoher Signifikanz ($> 8\sigma$).

- Bestimmung des differentiellen Energiespektrums von PG1553+113 im Energiebereich $80\,\mathrm{GeV} < E < 600\,\mathrm{GeV}$. Generelle Übereinstimmung mit den Resultaten einer unabhängigen Analyse [A$^+$06c] und mit Messungen der H.E.S.S. Teleskope [A$^+$06b].

Die Analyse - getestet an Monte Carlo Daten - bewies ihre Zuverlässigkeit bei der Untersuchung der Gammastrahlungsemission des Krebsnebels, wobei ein hochsignifikanter Exzess im Energiebereich unterhalb 100 GeV in nur 1.7 h nachgewiesen werden konnte.

Die Analyse von Daten des BL Lac Objekts PG1553+113 ergab signifikante Exzesse für Beobachtungen in den Jahren 2005 und 2006. Das kombinierte *alpha*-Histogramm zeigt ein Signal mit einer Signifikanz, die 8σ überschreitet.

Bei der weiteren Analyse konnte ein differentielles Energiespektrum für die kombinierten Daten aus den Jahren 2005 und 2006 erstellt werden (integraler Fluss oberhalb von 200 GeV: $F_{>200\,\mathrm{GeV}} = (1.7 \pm 0.3_{stat}) \cdot \frac{10^{-12}}{\mathrm{cm}^2\,\mathrm{s}}$, spektraler Index: $\Gamma = 3.6 \pm 0.3_{stat}$). Dieses Spektrum konnte daraufhin verwendet werden, um die (unbekannte) Rotverschiebung von PG1553+113 auf $z < 0.68$ (2σ Konfidenzniveau) einzuschränken. Mit einem einfachen SSC (synchrotron self-Compton) Modell konnte die elektromagnetische Emission von PG1553+113 vom optischen bis zum hochenergetischen VHE Energiebereich beschrieben werden. Wie aufgrund früherer Analysen anderer Blazare erwartet, konnte ein Verhältnis η zwischen e^--Energiedichte und magnetischer Energiedichte signifikant grösser als 1 abgeleitet werden.

Introduction

Within the new and rapidly developing field of astroparticle physics, the branch of γ-ray astronomy plays an important role in enlarging our understanding of the most extreme processes in the universe. A first motivation of γ-ray astronomy was a tracing of the sources of cosmic rays - a stream of charged, energetic particles impinging on the earth's atmosphere from all directions (see section 1.1). However, the investigation in particular of very active galaxy cores - the so-called active galactic nuclei (AGNs)- revealed extreme physics, most notable through a dramatic time variability of the observed γ-ray flux intensity in the order of days or even hours. Thus, γ-ray astronomy became a valuable tool for testing the physics, which is at work in the AGN 'machine' (see section 1.4).

As the γ-rays, emitted by the AGN jet, propagate through the extragalactic space, they are absorbed when interacting via pair production with the photon field of the extragalactic background light (EBL, see section 1.5). Observations within the so-called gap range of $50\,\text{GeV} - 300\,\text{GeV}$ (see section 3) are necessary for detections of AGN at high redshifts, since the EBL photons cause an energy cut-off limiting the observed γ-ray emission to energies mostly in the sub-TeV range for sources at redshifts $z \gtrsim 0.3$. Here the technical innovations of the MAGIC telescope, making an analysis at 'low' γ-ray energies possible, become important:

The MAGIC telescope was designed to lower the energy threshold of current generation imaging atmospheric Cherenkov telescopes. This goal is basically achieved by means of a large reflector together with a highly sensitive camera (see chapter 3), which both increase the telescope's sensitivity, so that an operation at low Cherenkov photon densities is possible. Furthermore, the parabolic overall reflector shape conserves the time profile of the Cherenkov light front coming from the shower, so that, along with the fast FADC system, the new possibility of an arrival time analysis arises.

The challenge of the data analysis at low energies is caused by the small Cherenkov light output of the shower, which results in small images, i.e. the image is built up by only a few pixels containing a moderate signal. An additional problem in the analysis is the presence of the unavoidable night sky background (NSB), which introduces noise to the image.

In order to meet this challenge, optimized methods of signal extraction and image cleaning were implemented, which try to extract as much signal from the shower as possible, while rejecting noise from the NSB, giving a sound basis for the subsequent analysis.

The main aim of this thesis is to provide advanced methods for the so-called g/h-separation, an analysis step, which is decisive in reducing the overwhelming background of hadronic showers in order to make a detection of weak γ-signals possible. In the presented work, an advanced method of g/h-separation using a tree classifier approach

based on the Random Forests algorithm is introduced. This method makes a simultaneous use of different image parameters possible, even if there are strong correlations between them.

In this thesis the need for an efficient analysis becomes evident in the investigation of VHE γ-emission from the BL Lac type AGN PG1553+113. This source, most probable located at a redshift z exceeding 0.3, shows merely a weak signal of about 4% Crab Nebula flux at 500 GeV (see chapter 7).

Contents

Chapter 1

High energy gamma-ray astronomy

1.1 Historical introduction

The history of the detection of cosmic rays is more than 100 years old. At the end of the 19th and the beginning of the 20th century Wilson and Elster & Geitel were dealing with measurements of the residual conductivity in air. In 1903 Rutherford found that a shielding around an electroscope reduced this residual conductivity, and it was near at hand to make the radioactivity recently discovered by Bequerel responsible for this effect. At first, the earth's natural radioactivity was a candidate for causing the conductivity of air [Gru00, Gru05].

Indeed, Wulf found out in 1910 that an electrometer lifted to the Eiffel tower showed a decreased intensity, which seemed to confirm the terrestrial origin of the ionizing radiation. But then the first measurements of Hess (1911/1912, Nobel price 1936) by means of balloon flights up to heights of 5 km showed that there was another component besides the terrestrial radiation, which increased with the altitude. This extra-terrestrial radiation was confirmed two years later by Kohlhörster, and soon new detection methods were found: Wilson developed 1912 a cloud chamber ('Wilson Chamber'), by which he made the tracks of charged particles visible. When Clay showed in 1927, that the intensity of the extra-terrestrial radiation depended on the geographical latitude, this was a clear hint to the charged particle nature of the so-called *cosmic rays* [Gru00, Gru05].

The investigation of the cosmic rays, which were a 'natural laboratory' for high energy physicists, providing energetic collisions beyond the reach of experimental techniques in the early era of high energy physics, developed quickly after the second world war. In particular, the subject of γ-ray astronomy was boosted by a seminal paper of Morrison (1958), where he pointed out the important role of γ-ray observations [Sch01]: The production mechanisms and emission regions of γ-rays are closely connected to the unknown cosmic ray sources. Since γ-rays, unlike charged particles, do not 'feel' magnetic fields and thus travel on straight trajectories, they provide valuable information about the long searched for sources of cosmic ray acceleration. Yet the expected γ-ray-fluxes were overestimated and it took some time until γ-ray sources could be detected. First detections in the high energy (HE) regime (about 100 MeV $-$ 10 GeV) were made by space borne detectors [Wee03]. The sensitivity of space borne experiments is limited due to their small area, which makes long observation times necessary, especially when observing in the high

HE and low VHE gamma regimes (VHE = very high energy , about 10 GeV up to several TeV). Therefore, ground based atmospheric Cherenkov telescopes were developed, which - by their working principle - have much larger effective collection areas (see chapter 3). A brief overview of the main historical 'mile stones' (detectors and their discoveries) is given in the following.

First detections of HE γ-rays from the space and the earth's atmosphere were made by the Explorer 11 experiment (1961). The Vela-Satellites discovered the γ-ray burst phenomenon (1967). In 1968 OSO-3 was able to detect HE γs from the Milky Way. The first point-like γ-source (the Geminga source) was found in 1972 by the SAS-2 satellite, which was later identified with a pulsar during the EGRET mission (see below). The first map of the Milky Way with an identification of 24 γ-sources was obtained by the COS-B mission in 1975.

In 1989 the Crab Nebula was detected by the Whipple Cherenkov telescope as the first source of TeV gamma rays (inside the so-called very high energy (VHE) range). The very successful Compton Gamma Ray Observatory (CGRO, operated from 1991 to June 2000), which consisted of four detectors (with the EGRET detector being sensitive in the 100 MeV $-$ 10 GeV energy range) covering the full HE region, produced all-sky maps of high resolution, revealing a wealth of galactic and extragalactic point sources, among them the blazar-type AGN Mkn421 (more precisely: a BL Lac object). Mkn421 was later (1992) detected by Whipple as a TeV source and soon confirmed by other experiments [Sch01]. Observation of similar BL Lac objects revealed the next TeV γ-ray source: Mkn501 (detected in 1995). This was the first detection of a source by a Cherenkov telescope, which had not been seen before by a space borne or any other detector.

Although Mkn501 was later on detected by EGRET during a flaring state, this VHE detection showed for the first time that observations by means of Cherenkov telescopes are a legitimate and powerful branch of astronomical investigations.

1.2   Spectrum and composition of Cosmic rays

The primary cosmic rays are investigated by means of space borne and balloon detectors (direct detection) as well as by air shower arrays (indirect detection, reconstruction of primary particles' properties detecting secondary air shower particles). In terms of numbers the composition of primary cosmic rays is (see [Per03] and [Lon92])

Protons:	86 %
α-Particles:	11 %
Heavy Nuclei:	1 %
Electrons:	2 %

The proportions of the major cosmic ray components given above are relatively constant with energy, all particles penetrating the geomagnetic field due to a high-enough rigidity are considered, more details can be found in [Gai90].

Furthermore, very small fractions of e^+, $\overline{p}$ are observed, which originate from interactions of the primary cosmic rays with interstellar gas. Also, neutral particles like ν, $\overline{\nu}$

(main sources are the sun and supernova explosions) and γ's (as diffuse galactic emission and emission from point sources, making up a fraction as small as $< 10^{-4}$) are present. The energy density of cosmic rays in the Milky Way is of the order $\sim 1\,\mathrm{eV\,cm^{-3}}$, which makes them a major contribution to the total energy density taking into account the energy densities of starlight ($\sim 0.6\,\mathrm{eV\,cm^{-3}}$), the cosmic microwave background ($\sim 0.26\,\mathrm{eV\,cm^{-3}}$) and galactic magnetic fields ($\sim 0.25\,\mathrm{eV\,cm^{-3}}$ arising from a magnetic field strength of $\sim 3\,\mu\mathrm{G}$).

The spectrum of the cosmic rays shows a typical power law shape with spectral index of -2.7 in the range from a few GeV up to $3 \cdot 10^{15}$ eV (the so called 'knee').

$$\frac{dN}{dE} \sim E^{-2.7} \tag{1.1}$$

Beyond $3 \cdot 10^{15}$ eV the spectrum steepens with a spectral index of -3. For energies exceeding $3 \cdot 10^{18}$ eV (the so called 'ankle') a flattening is observed (see figure 1.1).
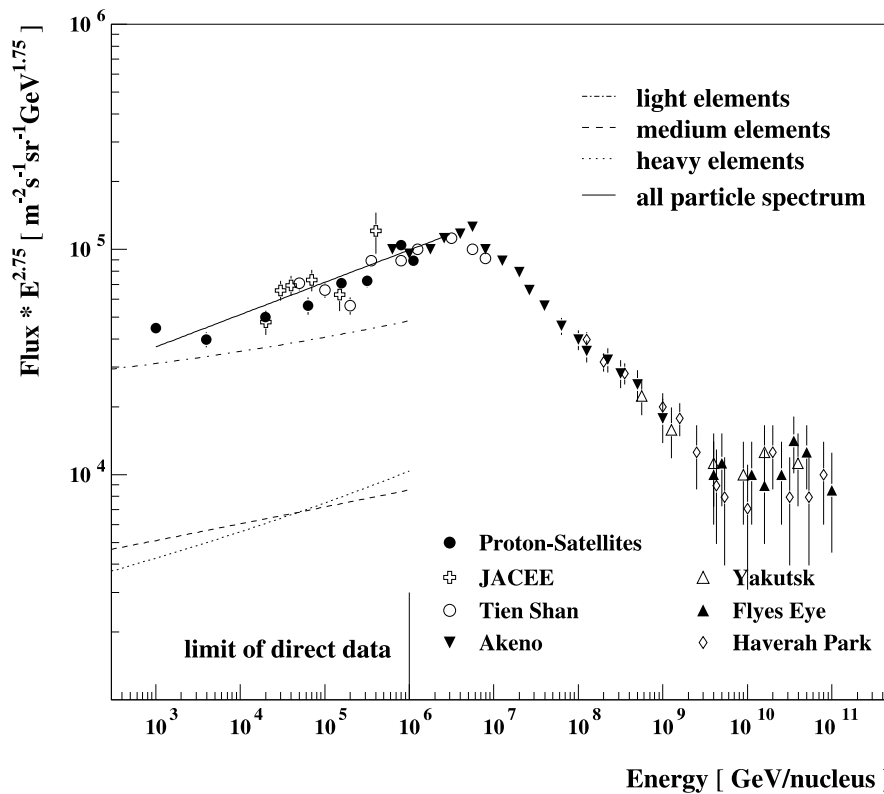
Figure 1.1: The all particle cosmic ray spectrum. In order to emphasize the changes in the spectral index, the flux is multiplied by $E^{2.75}$. The data were obtained from space borne and balloon experiments up to the limit of direct detection. For higher energies air shower arrays, having larger effective collection areas, are the main detectors.

Figure 1.2 shows the abundances of elements in cosmic rays in comparison with the solar element abundances. An overall similarity as well as differences for certain element groups are evident:

- Both solar and cosmic ray abundances show the odd/even effect:
  The nuclei with even Z/N are more strongly bound, hence they appear more frequently as end products of stellar fusion processes. In particular, the abundances in CNO and Fe are very similar, pointing to a stellar origin of most of the cosmic rays.

- Li, Be and B have low Coulomb barriers and, being weakly bound, their abundance in stars is very low. This element group appears more rich in cosmic rays due to spallation of carbon and oxygen in interstellar gas. One can derive a lifetime of $\sim 3 \cdot 10^6$ years from the Li-, Be- and B-abundance in cosmic rays and the known density of interstellar matter [Per03]. In an analogous way one can explain the more frequent abundance of Sc, Ti and Mn due to spallation of Fe and Ni.
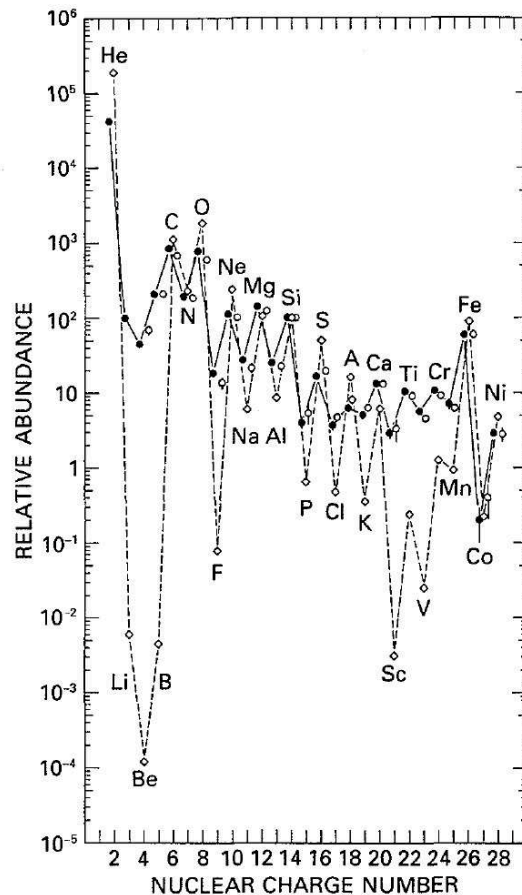
Figure 1.2: Abundances of elements in cosmic rays in comparison with the solar abundances. Figure taken from [Lon92].

Today, after 100 years of research, the origin of the cosmic rays is still basically unknown (see below for a discussion on potential sources). The open questions are:

- What are the sources of cosmic rays?

- What are the production and acceleration mechanisms?

Supernovae and Supernova remnants (SNRs) are the main candidates for sources of galactic cosmic rays of energies up to the knee region, yet the origin of the knee (change of acceleration mechanism?) is unknown. Particles with energies exceeding the ankle (so-called ultra high energy cosmic rays, shortly UHECR) are usually considered to be of extragalactic origin, since their Larmor radii in the galactic magnetic field are too large for a containment in our galaxy.

AGNs (Active Galactic Nuclei) are the most probable astrophysical sources of UHECR. They can accelerate particles up to 10^{21} eV, provide the needed UHECR emissivity and explain most naturally the observed small-scale anisotropies [BGG05]. There is also evidence for a directional correlation of UHE particles with BL Lacs [BGG05]. Another hint are the so-called 'orphan flares', i.e. γ-ray flares without simultaneous X-ray flare, which may indicate hadronic acceleration (see e.g. [K$^+$03]).

As already mentioned above, HE and VHE γ-emission is closely connected to the sources of the cosmic rays. Thus, the first motivation of γ-ray astronomy was a tracing of the CR sources. Nevertheless, since observations of AGN revealed extreme physics (short time variabilities, beaming effects, see section 1.4), a shift of interest from a mere search of CR sources to the objects themselves with their inherent physics took place.

In the following a description of γ-ray production processes and possible sources, with a focus on Active Galactic Nuclei, is given.

1.3 Production processes and Sources of gamma-rays

In classical astronomy, particularly in optical observations, the universe turns out to be dominated by thermal radiation showing the well-known black body spectrum. Yet, already in the case of 1 MeV gamma-rays one would need temperatures in the order of $2 \cdot 10^9$ K (for comparison: the sun shows a core temperature in the order of 10^7 K) to explain their emission as thermal process demanding conditions, which could be met only in extreme fireballs (e.g. GRBs, see below). Thus, the emission of gamma-rays is dominated by non-thermal processes. The following overviews, showing the relevant production processes as well as the corresponding possible sources of γ-rays, have been adopted from [Sch01], [Wee03] and [MW86]. For the most important gamma-ray production mechanisms of synchrotron radiation and inverse Compton scattering the spectral indices of the expected power law gamma energy distribution are given in terms of the parent electron power law energy distribution, which is assumed to be $I_e(E_e) \sim E_e^{-\Gamma}$.

- *Charged particles in strong electric or magnetic fields*

    - *Bremsstrahlung*
      If an electron passes closely by an atomic nucleus, it will be deflected in its electric field. The acceleration (or deceleration) causes a radiative energy loss of the electron, the so-called bremsstrahlung.

      The typical spectrum of the bremsstrahlung photons in the non-relativistic case is $\frac{dN_\gamma}{dE_\gamma} \sim \frac{1}{E_\gamma}$.

For higher energies (ultra-relativistic electrons), the spectrum shows a power-law shape with the a spectral index identical to that of the parent electron distribution [MW86], i.e. $\frac{dN_\gamma}{dE_\gamma} \sim E_\gamma^{-\Gamma}$

Critical parameters in the production of bremsstrahlung are the differential electron spectrum I_e (which determines the bremsstrahlung spectral index, see above) and the target gas density n_g.

– *Synchrotron radiation*
In the early era of radio astronomy polarized emission from radio galaxies and SNRs (super nova remnants) was detected, which could be successfully explained as synchrotron radiation caused by (ultra-) relativistic electrons moving in magnetic fields. Synchrotron radiation is known from accelerator experiments, where less energetic particles but stronger magnetic fields are involved (magnetic fields in cosmic sources are generally much weaker). In a synchroton process an electron is moving in a magnetic field, therefore being accelerated, and radiates like a dipole with the so-called gyration or Larmor frequency $\omega_L = \frac{eB}{m_e}$. At relativistic energies the emission is strongly beamed into a cone of angle $\theta \approx \frac{m_e c^2}{E}$. The maximum power is emitted at $\omega_c = \frac{3}{2}\omega_L \gamma^2 \sin\phi$ with ϕ being the angle between the particle trajectory and the direction of the magnetic field. Thus, a relativistic electron in a typical interstellar magnetic field emits synchrotron radiation at [Sch01]

$$E_\gamma \cong 0.05 \left(\frac{E_e}{\text{TeV}} \right)^2 \frac{B}{3\,\mu\text{G}} \text{ eV} \qquad (1.2)$$

For a power law electron spectrum as introduced above the synchrotron photon spectrum becomes [MW86]

$$\frac{dN_\gamma}{dE_\gamma} \sim E_\gamma^{-\frac{\Gamma+1}{2}} \qquad (1.3)$$

HE synchroton emission can only originate in regions of extremely high magnetic fields, e.g. close to a neutron star's surface (here magnetic field strengths exceeding 10^{10} G can be reached). In such environments also curvature radiation becomes important (see below).

– *Curvature radiation*
Magnetic field lines in the vicinity of neutron stars can provide a 'bending' acceleration to charged particles, which is sufficient for electrons to produce 'curvature γ-rays'. Here, the acceleration is due to charged particles following the magnetic field lines' curvature, whereas in a synchrotron process the acceleration is due to the particle 'spinning' around the magnetic field vector direction.

● *Inverse Compton Scattering*
The up-scattering of low energy photons by collisions with high energy electrons is called inverse Compton scattering (since it describes the inverted Compton scattering process). This process is important in regions of comparably high (soft) photon

densities n_{ph} and/or a sufficient population of relativistic electrons, quantified by the differential energy spectrum I_e (VHE emission of AGNs (see section 1.4) and SNRs).

Mean gamma energies are calculated for electron energies in the Thomson regime and in the (relativistic) Klein-Nishina regime:

$$\langle E_\gamma \rangle = \begin{cases} \frac{4}{3}\gamma_e^2 \cdot < \varepsilon > & \text{for } \gamma_e\varepsilon \ll m_ec^2 \text{ (Thomson limit)} \\ \frac{1}{2} < E_e > & \text{for } \gamma_e\varepsilon \gg m_ec^2 \text{ (Klein-Nishina limit)} \end{cases} \tag{1.4}$$

$$\frac{dN_\gamma}{dE_\gamma} \sim \begin{cases} E_\gamma^{-\frac{\Gamma+1}{2}} & \text{for } \gamma_e\varepsilon \ll m_ec^2 \\ E_\gamma^{-\Gamma} & \text{for } \gamma_e\varepsilon \gg m_ec^2 \end{cases} \tag{1.5}$$

Here, E_e denotes the parent electron's energy and ε is the seed photon's energy.

A useful approximate equation for the gamma energy in the Thomson regime is given by:

$$E_\gamma \cong 1.3 \left(\frac{E_e}{\text{TeV}}\right)^2 \left(\frac{E_{ph}}{2 \cdot 10^{-4}\,\text{eV}}\right)\,\text{GeV} \tag{1.6}$$

- *Nuclear transitions*
  Nuclear transitions can be seen as HE line emission. Characteristic energy levels of importance for gamma-astronomy are first-excited states of the nuclei C^{12} (4.438 MeV), O^{16} (6.129 MeV) and Mg^{26} (1.809 MeV). The excitation is taking place in regions of nucleosynthesis or cosmic ray - matter interaction.

- *Decay and Annihilation*
  In collisions of cosmic ray protons with ambient-gas nuclei (mostly hydrogen gas) the lightest mesons (pions) are likely to be produced, among these also the neutral π^0. This short-lived meson decays into two gammas with characteristic energies of $E_\gamma = 70$ MeV in the rest frame of the pion.

  Another example for γ-ray production is the annihilation of particle anti-particle pairs, in particular electron positron pairs, which is observed as 511 keV annihilation line. When positronium is formed (a bound e^+e^- system similar to hydrogen, but also to helium since triplet and singulett states exist with antiparallel and parallel spins respectively), a decay into 3 photons sharing 1.022 MeV (and peaking at 511 keV with a continuum towards lower energies) takes place.

- *Charged particles bound in strong magnetic fields*
  Analogously to the quantized energy levels in atoms, 'Landau' energy levels $E = \hbar\omega_L$ with the cyclotron frequency $\omega_L = \frac{ZeB}{\gamma m_0}$ can be formed in strong magnetic fields. In the vicinity of highly magnetized neutron stars the corresponding cyclotron lines are within the X-ray regime.

Absorption and interaction of γ-rays with matter is related to the 'famous' three possible processes: Photo absorption (dominant below 100 keV), Compton scattering (dominant for medium energies of 100 keV − 5 MeV) and pair production (threshold energy

1.022 MeV neglecting the atom's recoil). It should be mentioned that pair production can not only occur in the vicinity of atomic Coulomb fields, but also within low energy background fields (cosmic microwave (CMB) and infrared background (CIB), see section 1.5) leading to an opaqueness to VHE γ-rays, the so-called gamma-ray horizon.

In the listing below, possible γ-ray sources are given and briefly explained [Sch01]:

- *Fireballs*: γ-ray bursts (thermal γ's possible!)

- *Explosive events of extreme energy density*: supernovae and novae

- *Energetic collisions*:

    - Accreting compact objects often found as binary systems
    - Particle jet sources like Microquasars, AGNs ($\rightarrow$ Compton upscattered gammas, π^0-production)
    - Cosmic Ray collisions with ambient matter
    - Solar flares (accounting for the lower γ-energy range up to a few GeV)

- *Charged-particle beams in the vicinity of compact objects (strong gravitational and magnetic fields)*: AGNs ($\rightarrow$ synchrotron radiation)

A more detailed list of observational objects with an emphasis on the detection potential of MAGIC is given in chapter 3.

1.4   Active Galactic Nuclei (AGNs)

About 3 % of all galaxies are 'active' [Sch01]. In contrast to 'normal' galaxies, which are basically an assembly of stars, they show highly luminous nuclei (hence the term Active Galactic *Nucleus*) active in various wavelength regions. Further differences between AGNs and normal galaxies are:

AGNs show broadened emission lines; the luminosity is peaking in the IR, UV, X-ray or γ-ray region. On the other hand, normal galaxies show absorption lines and the luminosity is peaking in the optical, being the sum of the star light. AGNs are furthermore often dominated by non-thermal emission, they show a high variability on short time scales of the order of days or even hours.

Figure 1.3 shows a skymap in gamma-rays as observed with the EGRET satellite experiment. Many of the point sources at higher galactical latitudes are AGN.

The source of energy in AGNs is believed to be the gravitational sink created by a very massive core: Gravitational energy is converted into radiation and kinetic energy via accretion of matter onto a central supermassive black hole, which leads to the formation of an accretion disk due to conservation of the inflowing matter's angular momentum. By means of simple considerations one can derive constraints on some physical parameters:

- The possible size r of the emission region is constrained by light travel arguments taking into account the observed high variability on time scales Δt:
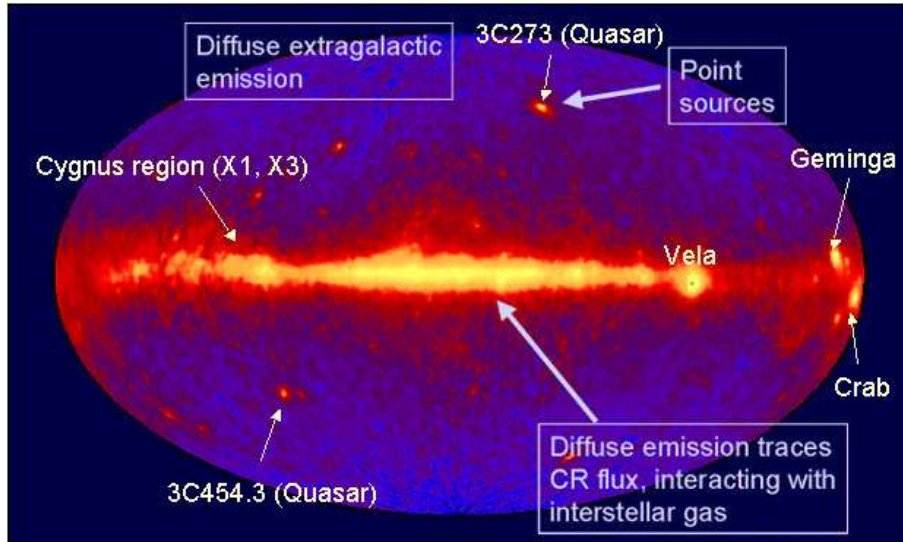
Figure 1.3: Skymap in galactical coordinates for $E_\gamma > 100\,\text{MeV}$ (3rd EGRET catalogue, see [H$^+$99]). Many of the point sources at higher latitudes are AGN. Adopted from [HEA06].

$$r < \frac{c\Delta t}{\delta} \tag{1.7}$$

Here $\delta = 1 + z$ denotes the redshift z related Doppler factor, c is the speed of light. A variability of ~ 1 day (as observed for many Quasars) and a redshift of $z \sim 1$ yields a region of the size identical to our solar system, where the variable flux is produced; hence the term Active Galactic **Nucleus** is justified.

- The produced luminosity can be set in relation to the accretion rate in the following way [Sch01], with ε being the energy conversion efficiency and m_{acc} the accreted matter:

$$L_{acc} = \varepsilon\,\dot{m}_{acc}c^2 \tag{1.8}$$

Here, $\dot{m}_{acc}$ is the accretion rate. With $\varepsilon \sim 10\,\%$ the energy conversion efficiency in an accretion process is much larger than what is known from nuclear fusion, where values of $\varepsilon \sim 0.7\,\%$ are reached. There exists an upper limit for the possible accretion powered luminosity, which arises from the equilibrium condition between gravitational force and radiation pressure (a too big accretion rate would lead to a dominating radiation pressure, which then stops the accretion process). The so called Eddington limit, at which the accretion disk is still dynamically stable, can be expressed for a spherical accretion process and Compton scattering in the Thomson regime as a function of the central mass M in the following way [Sch01]:

$$L_{Edd} = \frac{4\pi G m_p c}{\sigma_T} M \qquad (1.9)$$

Here G is the gravitational constant, m_p denotes the proton mass and σ_T stands for the Thomson cross section. Putting in numbers, with $M_\odot$ being the solar mass, one yields:

$$L_{Edd} = 1.3 \cdot 10^{38} \frac{M}{M_\odot} \text{ erg s}^{-1} \qquad (1.10)$$

From observed luminosities, which are ranging from $(10^{12} - 10^{13})L_\odot$ with $L_\odot$ as solar luminosity, it follows that the supermassive black hole exhibits a mass of the order $M = (10^6 - 10^{10})M_\odot$.

The different existing types of AGNs can be classified according to the scheme shown in table 1.1. Here the optical emission line properties (AGN types 0, 1 and 2) are set in relation to a decreasing angle between the jet axis and the line of sight. The reason for a distinction between radio loud and radio quiet AGNs is basically unknown; the assumption is, that in the case of radio quiet AGNs the jet is developed just weakly. The Blazar sub-class of AGNs, which exhibit a significant gamma-ray emission, is made up by BL Lac objects (named after the prototype galaxy BL Lacertae) and Flat Spectrum Radio Quasars (FSRQs). These objects are characterized by a strong variability of their emission, a flat radio spectrum and optical polarization. Blazars are radio loud AGNs observed at very small jet-to-line-of-sight angles.

One has to keep in mind that the grouping of this scheme (of any scheme) is not strict: There is an assumption that the hard distinctions between Fanaroff Riley (FR) I and II galaxies (exhibiting differently strong developed radio jets) is in fact a smooth transition, which appears strong only due to a limited sample of observations [Urr03]. Furthermore, Quasi Stellar Objects (QSOs) may be just distant Seyfert galaxies, where only the core remains visible (Seyferts are spiral galaxies, whereas radio-loud AGNs are mostly elliptical galaxies).

Figures 1.4 and 1.5 depict the corresponding so-called unified AGN model:

A central supermassive black hole of mass $10^6 - 10^{10}$ $M_\odot$ is sourrounded by a highly relativistic, rotating accretion disc of plasma. The plasma is formed since accreted gas heats up due to friction and starts to emit thermal UV and soft X-rays. This radiation excites atoms and molecules in nearby gas clouds. When the clouds are located close to the accretion region, they show Doppler broadened optical emission lines; further distant and therefore slower moving clouds show narrow lines. Two jets are emitted probably close to the poles of the black hole. Charged particles within these jets can be accelerated to very high energies by means of shock fronts and produce radiation via synchrotron and inverse Compton scattering processes. Synchrotron photons are produced at different jet-core distances yielding radiation with wavelengths from X-ray over optical down to the radio range. High energy photons are produced as synchrotron and, more energetic, inverse Compton radiation. There is also an obscuring torus (cold gas and dust) at some
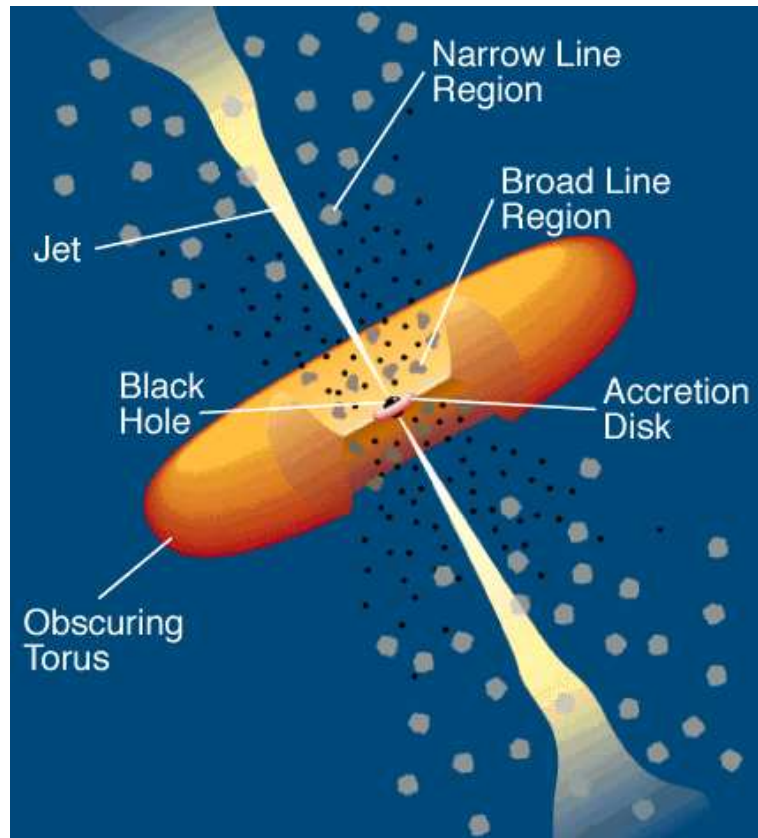
Figure 1.4: Unified AGN model, taken from [UP95]: The basic AGN setup is depicted. A central supermassive black hole is surrounded by an accretion disk of hot plasma emitting a thermal spectrum of UV and soft X-ray photons. Further outside an obscuring torus of cold molecular matter extends from an inner diameter of 1.5 pc up to an outer diameter of 30 pc [Sch01]. Narrow and broad line emission regions correspond to gas clouds moving at velocities $v < 2000\,\mathrm{km\,s^{-1}}$ and $v \geq 2000\,\mathrm{km\,s^{-1}}$ respectively. Thus, the broad emission lines are caused by a stronger Doppler broadening (the line broadening is a result of the fast movement of many small gas clouds, it is not a thermal broadening). The black dots inside the broad line - and reaching up to the narrow line region symbolize a hot electron corona. Scattering at this corona probably makes a weak, broad emission line observation at higher jet-to-line-of-sight angles possible [Sch01].

Radio loudness	Optical emission line properties			
	Type 2 *(narrow lines)*	*Type 1* *(broad lines)*	*Type 0* *(weak/absent)*	
Radio-quiet *(85-90 %)*	Seyfert 2	Seyfert 1 QSO		↑ ?
Radio-loud *(10-15 %)*	NLRG = $\begin{cases} \text{FR I} \\ \text{FR II} \end{cases}$	BLRG SSRQ, FSRQ	Blazars = $\begin{cases} \text{BL Lac} \\ \text{FSRQ} \end{cases}$	↓ ?
	Decreasing angle between jet axis and line of sight →			

Table 1.1: AGN classification adopted from [Sch01] and [UP95]. The reason for the appearance of jets (radio-loud AGN) is basically unknown, among possible distinctive physical parameters the black hole spin is under consideration [UP95]. The abbreviations are explained in the text.

distance from the center, which inhibits a direct observation of the accretion region at large jet-to-line-of-sight angles.

The different properties of observed AGNs are now explained within this model according to figure 1.5. For large angles between the jet axis and the line of sight a radio galaxy (NLRG, type RF I or II) or a Seyfert 2 galaxy is observed. Only the narrow optical emission lines can be seen since the dust torus is obscuring the broad line region, which is located closer to the black hole. For a smaller jet-'line of sight' angle also these lines become visible and a Seyfert 1, QSO (possibly just located further away than Seyfert 1) or a BLRG (Broad Line radio Galaxy) with steep (SSRQ) or flat (FSRQ) radio spectrum is observed. The distinction between radio-loud and radio-quiet AGN demands no or just a weakly developed jet in the lower part of figure 1.5. Furthermore, the difference between FR type I and II AGN is probably explained by differently strong developed jets [Urr03]. Within this model it is assumed, that FR I type AGNs correspond to BL Lac type blazars and the FR II type AGN are identical to FSRQ type blazars (from the physics point of view, just the observation angle differs). Both types of blazars show just weakly developed optical lines or no optical emission lines at all. Here the observer directly looks 'down the gun barrel' into the jet and relativistic beaming effects become important.

As already mentioned, strongly collimated highly relativistic jets or blobs are emitted perpendicular to the accretion disk, probably from the poles of the black hole. Since the jets are highly relativistic, typical properties of the emitted γ-rays (γ-production mechanisms are discussed below) can be observed:

- *Apparent superluminal motion*
  In radio images of Blazars superluminal motion of gas blobs can be observed. This
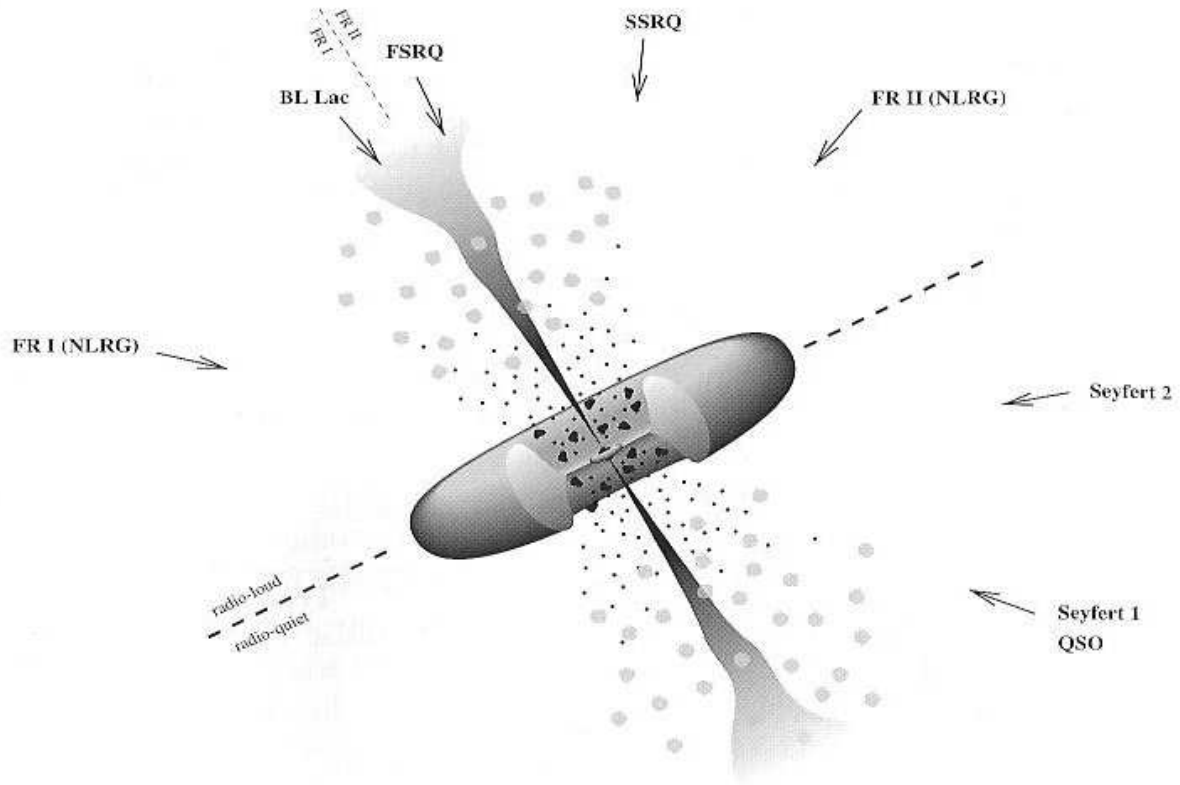
Figure 1.5: Unified AGN model, taken from [Sch01]: The orientation dependent classification is pointed out. The distinction between radio-loud and radio-quiet AGN requires a weak jet or no jet at all in the lower part of the figure. FR I and II type AGN corresponding to the left and right upper part of this figure are probably distinguished from each other by jets with different strength (FR I AGNs showing a weaker jet).

geometrical effect is explained by a highly relativistic jet having a small angle with the line of sight.

- *Observation of HE and VHE gamma-rays and size of emission region*
  Due to the observed strong flux variability on time scales as small as days or even hours the size of the emission region is limited (see equation 1.7) to a diameter of a few light-minutes or hours. For the observed luminosities of $L = 10^{48} - 10^{49}$ erg s^{-1} a problem due to the absorption of gammas though pair production arises:

$$\gamma + \gamma \rightarrow e^+ + e^- \tag{1.11}$$

The attenuation factor τ is related to the gamma density n_γ, the cross section $\sigma_{\gamma\gamma}$ for $\gamma\gamma$-pair production and the size R of the emission region according to:

$$\tau_{\gamma\gamma} = n_\gamma \sigma_{\gamma\gamma} R \tag{1.12}$$

The condition for opaqueness reads

$$\tau_{\gamma\gamma} \approx \left(\frac{L_\gamma}{4\pi R^2 m_e c^3} \right) \sigma_T R = \frac{\sigma_T}{4\pi m_e c^3} \frac{L_\gamma}{R} = 20 \left(\frac{L_\gamma}{10^{48} \text{ erg s}^{-1}} \right) \left(\frac{R}{10^{15} \text{ cm}} \right)^{-1} \geq 1 \tag{1.13}$$

This condition for opaqueness can be solved for an isotropic luminosity L_γ and the mentioned size of emission region yielding $\tau \approx 20$, which means that no significant gamma flux should be observed. This contradiction can be resolved if one assumes beamed, relativistic emission (which is also implied by the observed superluminal motion). In this case a Doppler boosting with a factor δ must be taken into account:

$$\delta = \frac{1}{\Gamma(1 - \beta \cos\theta)} \, , \; \beta = \frac{v_{jet}}{c} \, , \; \Gamma = \frac{1}{\sqrt{1 - \beta^2}} \tag{1.14}$$

The observed luminosity L_γ is now equal to the Doppler boosted intrinsic luminosity L'_γ

$$L_\gamma = \delta^n \cdot L'_\gamma \tag{1.15}$$

With Doppler factors of the order of 10 and values $n = 3 - 4$ (for current emission models [Sch01]) the intrinsic luminosity is a factor of 1000 or 10000 smaller, which now means that τ is uncritically around 0.1.

Figure 1.6 impressively shows the image of a radio galaxy. The jets as well as the dust torus are clearly seen supporting the unified AGN model.

As for the appearance of these jets, different models exist [Sch02]: Jets are a natural consequence of a rotating plasma disc in the presence of a magnetic field (magnetically driven and collimated jet) especially when a fast rotating so-called Kerr black hole is present (frame dragging in the black hole's vicinity guides the plasma closer to the event
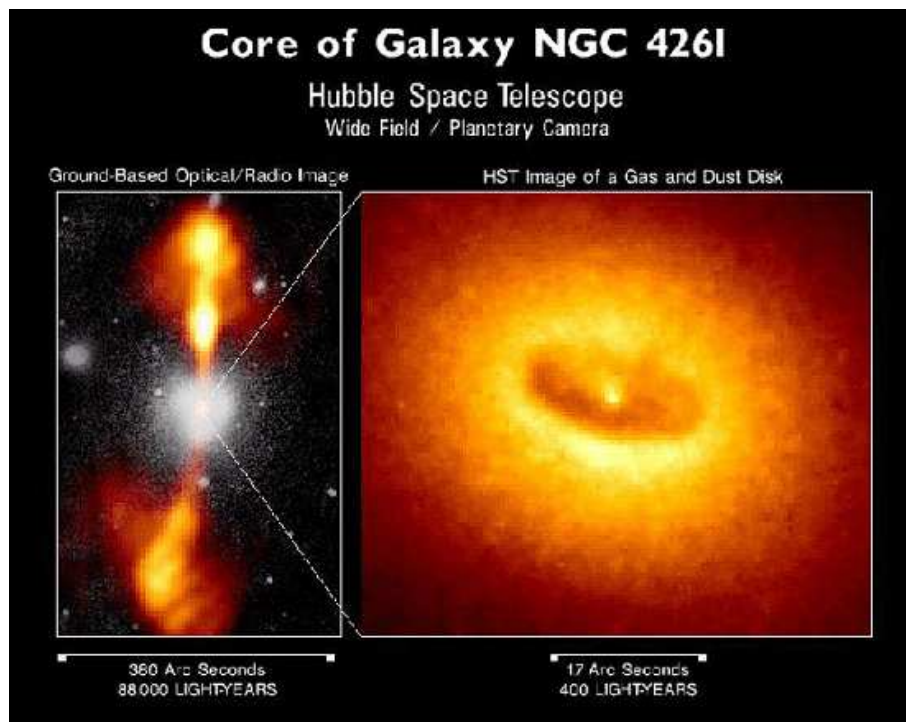
Figure 1.6: An overlay of optical and radio images (on the left) and a high resolution optical image from HST (on the right) showing the radio galaxy NGC 4261. The radio jets as well as the obscuring torus are clearly visible. Image taken from [STS06].

horizon and deeper into the gravitational well, so that more energy is available for a 'jet production'). These jets are likely to accelerate light particles. At present, the above mentioned magnetohydrodynamic production and collimation of jets is the leading model [MKU01].

Another possibility is a black hole acting as dynamo in the magnetic field of the accreted plasma. Now the event horizon behaves like a conductor and the black hole charges up at its poles. Here, within the strong electric field, $e^+ e^-$-pairs are created, and the equally charged particles (with respect to the static charge) are repelled (electrically driven, magnetically collimated jet).

Inside the jet particles can be accelerated via shock waves to very high gamma energies. The expected mean ratio of the final and the initial energy $\langle E_f/E_i \rangle$ is deduced from the so-called Fermi acceleration of first order (the name arises because the energy gain is proportional to the shock front velocity v). With $\beta = \frac{v}{c}$ and c being the speed of light, the average of the energy ratio reads [Gal02]:

$$\left\langle \frac{E_f}{E_i} \right\rangle \approx 1 + \frac{4}{3}\beta \qquad (1.16)$$

Assuming a fixed escape probability yields a power law spectrum for the energy of the accelerated particle:

$$\frac{dN}{dE} \sim E^{-\alpha}, \text{ with } \alpha = 2 \qquad (1.17)$$

In a relativistic approach a spectral index of $\alpha = 2.2 - 2.3$ is obtained. The maximum particle energy is limited due to [Gal02]:

- The maximum Larmor radius, which is determined by the size of the emission or shock region R_{sh} and the magnetic field strength B. The maximum energy can be expressed as

$$E_{max} \approx qB\gamma_{sh}R_{sh} \qquad (1.18)$$

  with γ_{sh} being the Lorentz factor of the shock region and q being the particle charge.

- Cooling down of electrons by radiation processes (synchroton, IC), which reduces the maximally achievable energy.

As for the nature of the accelerated particles, leptonic and hadronic models exist, which both predict a VHE gamma-ray production:

- *Leptonic models* (see figure 1.7)
  The VHE gamma-emission as a result of inverse Compton processes can be understood as SSC (synchrotron self Compton) or EIC (external inverse Compton) mechanism depending on the type of the soft seed photons. Within the SSC model the upscattering electrons themselves produce the soft X-ray photons in a preceding synchrotron process. In the EIC model the seed photons originate from the thermal X-ray radiation, which stems from the accreted plasma and may be scattered before the upscattering process.
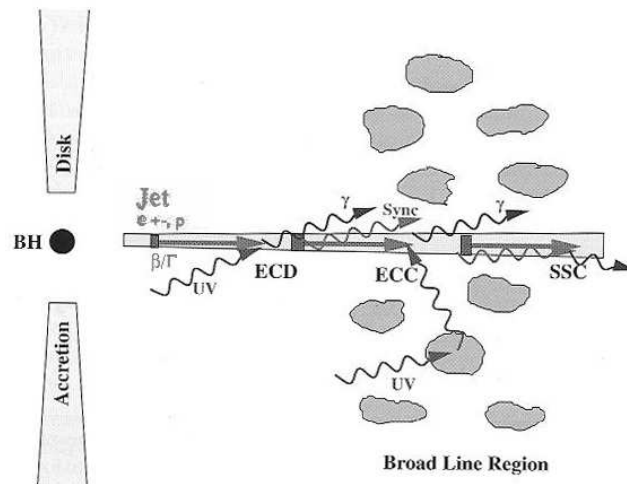
Figure 1.7: The VHE emission of AGN [Sch01]: In this model the γ-ray emission is produced inside the inner jet region, in a few 100 R_s distance from the core, where blobs of relativistic particles are accelerated in the jet (R_s = Schwarzschild-radius). BH = black hole, UV = UV-radiation from the accretion disk, Sync = synchroton emission from the jet. The different inverse Compton models involve: ECD = External Compton Scattering of accretion disk photons, ECC = External Compton Scattering of scattered accretion disk photons, SSC = Synchroton Self-Compton Scattering (here, electrons produce in a synchrotron process the seed photons for a subsequent inverse Compton scattering).
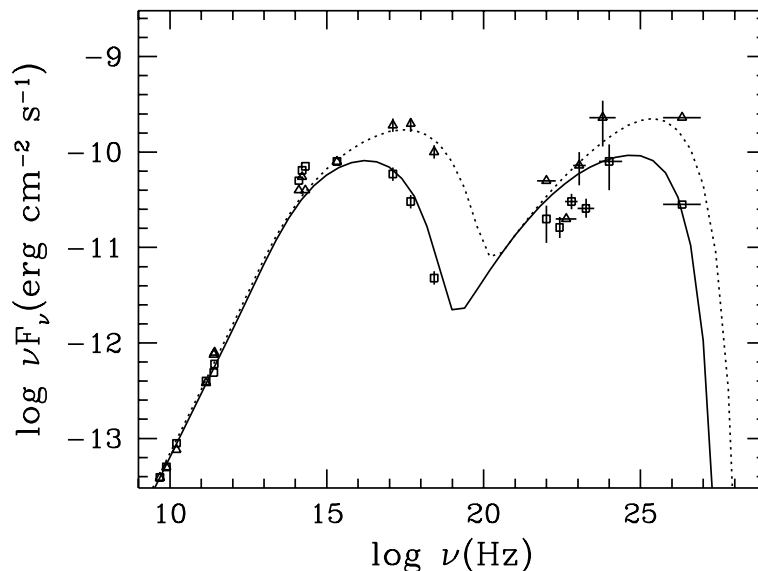
Figure 1.8: Double peaked spectral energy distribution (SED) of Mkn421 for high (triangles, dotted curve) and low state (rectangles, solid curve) taken from [Sch02]. The multiwavelength spectrum was fitted with a SSC model assuming a laminar emission region geometry.

- *Hadronic models*
  It may be that hadrons carry the bulk of the energy inside the jet. Photo-pion production sets in at an energy threshold of 10^{20} eV for protons. Since protons do not show strong radiation losses, they can be accelerated to these energies, and the following reactions take place:

  Production of neutral pions followed by their decay into gammas:

$$\begin{aligned} p + \gamma & \rightarrow & \pi^0 + \dots \end{aligned} \qquad (1.19)$$
$$\begin{aligned} \pi^0 & \rightarrow & \gamma + \gamma \end{aligned} \qquad (1.20)$$

  Production of charged pions and succeeding decay to muons, electrons and neutrinos:

$$p + \gamma \rightarrow \pi^{\pm} + \dots \qquad (1.21)$$
$$\pi^+ \rightarrow \mu^+ + \nu_\mu \qquad (1.22)$$
$$\pi^- \rightarrow \mu^- + \overline{\nu}_\mu \qquad (1.23)$$
$$\mu^+ \rightarrow e^+ + \overline{\nu}_\mu + \nu_e \qquad (1.24)$$
$$\mu^- \rightarrow e^- + \nu_\mu + \overline{\nu}_e \qquad (1.25)$$

  Production of electrons and positrons:

$$p + \gamma \rightarrow e^- + e^+ + \dots \qquad (1.26)$$

It shall be pointed out that the gamma-emission in the hadronic models is also produced by the inverse Compton process involving energetic leptons. The high optical depth for pair production leads to successive pair production and inverse Compton scattering processes until the γ-energy falls below the threshold energy, for which the optical depth becomes smaller than 1. Then the γ-rays can escape.

Since in the above mentioned processes also neutrinos are produced, a signature caught by neutrino telescopes would immediately confirm the hadronic model. Up to now, the Leptonic model is more favored since the predicted shapes of synchroton and inverse Compton peaks match well with the observed spectra (see figure 1.8).

1.5   Absorption of gamma-rays

1.5.1   The extragalactic photon background

The fate of gamma-rays from their emission up to their detection depends on the traversed intergalactic medium, which is radiation dominated. The main components of this *Extragalactic Background Light (EBL)* are:

- *The Cosmic Microwave Background (CMB)*
  The residual radiation of the 'Big Bang' originates from a time, when the universe had cooled down to a temperature where radiation decoupled from matter (thus

the universe became transparent, about $4 \cdot 10^5$ years after the 'Big Bang'). The radiation at that moment of time was black body radiation (due to the equilibrium between radiation and matter), which afterwards continuously cooled down until to a temperature of $2.75\,\mathrm{K}$ at present. The corresponding spectrum shows high intensities at wavelengths $\lambda \gtrsim 300\,\mu\mathrm{m}$ (compare figure 1.9).

- *The Infrared and Optical Background*
  Red shifted starlight from all eras of the universe makes up the so called Cosmic Infrared and Optical Background extending from the far infrared ($100\,\mu\mathrm{m}$) up to the UV region ($0.1\,\mu\mathrm{m}$). Star formation is expected to be a major contribution to the EBL yielding the typical double peak structure [Wee03] (see figure 1.9): A direct contribution of star formation at short wavelengths ($1 - 15\,\mu\mathrm{m}$) and a contribution due to dust absorption and re-emission at longer wavelengths ($100 - 200\,\mu\mathrm{m}$). In particular the star formation era in the early universe may have a significant influence on the density of the EBL. This means vice versa, that a proper knowledge of the EBL density possibly yields information about this early star formation process.

  A direct measurement of the EBL is difficult due to foreground (galactic) infrared light caused by dust reflecting the sunlight and by starlight. So, in the region between a few $\mu\mathrm{m}$ and $100\,\mu\mathrm{m}$ there are no direct measurements, just upper and lower limits. The lower limits are derived from galaxy counts and are hard limits.

VHE gamma-rays emitted e.g. from a blazar have to cross this medium being attenuated by the following pair production process:

$$\gamma_{VHE} + \gamma_{EBL} \rightarrow e^+ + e^- \tag{1.27}$$

The corresponding cross section $\sigma_{\gamma\gamma}$ is peaking shortly after exceeding the threshold energy. The latter is given by the condition, that the center of mass energy $\sqrt{s}$ is equal to $2m_e c^2$

In the following equations E denotes the energy of the VHE-gamma whereas ε denotes the EBL-photon energy.

$$s = c^2 p_\alpha p^\alpha, \ p_\alpha = \begin{pmatrix} (E+\varepsilon)/c \\ \boldsymbol{p}_{VHE} + \boldsymbol{p}_{EBL} \end{pmatrix} \tag{1.28}$$

$$\Rightarrow s = 2E\varepsilon(1 - \cos\theta) \tag{1.29}$$

Here, θ denotes the angle between the VHE photon and the low-energy target photon (for head-on collision: $\theta = 180°$). Setting c to 1 for simplicity reasons, the threshold condition $s > (2m_e)^2$ leads to

$$E\varepsilon(1 - \cos\theta) > 2m_e^2 \tag{1.30}$$

$$\Rightarrow E\varepsilon x > m_e^2 \tag{1.31}$$

Here, in the last step $x = \frac{1 - \cos\theta}{2}$ was introduced.
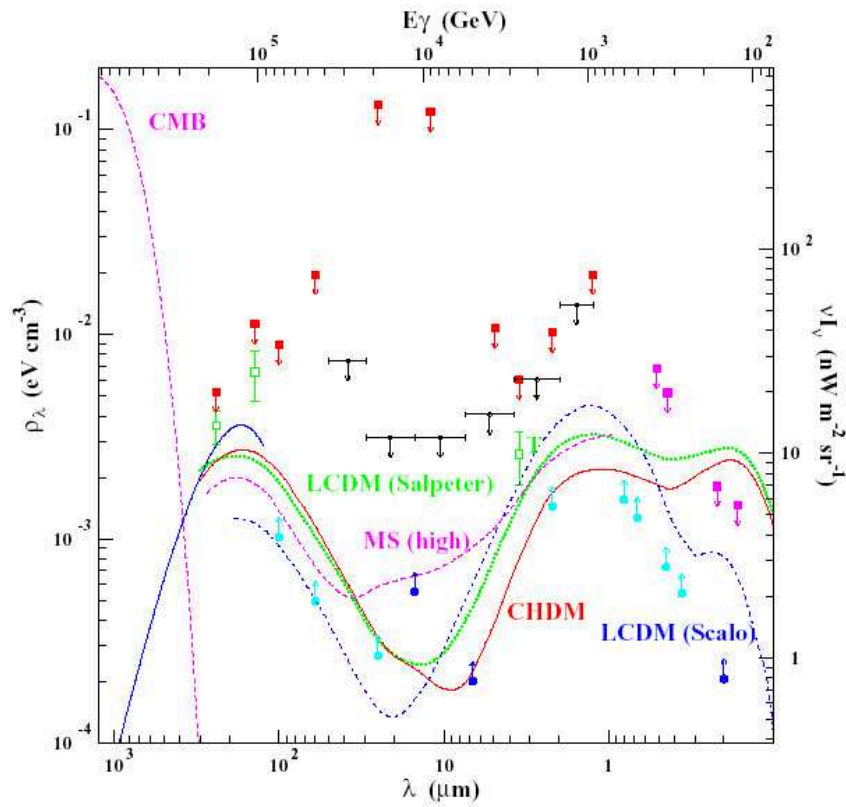
Figure 1.9: Extragalactic Background Light, measurements and models, taken from [Vas99a]. Note that the upper horizontal axis shows the energy of those VHE gammas, which are most effectively absorbed by EBL photons with wavelength given by the lower horizontal axis.

1.5.2 The photon-photon pair production cross section

In preparation of the analysis of data taken on the BL Lac PG1553+113 (see chapter 7), a detailed discussion of the photon-photon absorption according to the approach of Vassiliev [Vas99b] is given in the following. Recent developments in the astrophysical research make modifications necessary, which will be addressed thereafter. Introducing

$$q = \frac{m_e^2}{E\varepsilon x} \tag{1.32}$$

the threshold condition can be rewritten as $q < 1$. For fixed VHE-photon energy q runs from 1 to 0, when ε runs from the threshold value $\varepsilon = \frac{m_e^2}{Ex}$ to infinity. The cross-section of photon-photon pair production can now be expressed as a function of q [Vas99b].

$$\sigma(q) = \frac{3}{8}\sigma_T f(q) \tag{1.33}$$

$$f(q) = q\left[(1 + q - \frac{q^2}{2})\ln\left(\frac{1 + \sqrt{1-q}}{1 - \sqrt{1-q}}\right) - (1+q)\sqrt{1-q}\right] \tag{1.34}$$

$$q = \frac{m_e^2}{E\varepsilon x} \tag{1.35}$$

$$x = \frac{1 - \cos\theta}{2}, \ dx = \frac{\sin\theta}{2}d\theta \tag{1.36}$$

σ_T denotes the Thomson cross section. The dependence of $\sigma_{\gamma\gamma}$ on q - described by the function $f(q)$ in equation 1.34 - is shown in figure 1.10. For head-on collisions, i.e. $x = 1$, q simplifies to $q = \frac{m_e^2}{E\varepsilon}$, which can be rearranged to give an expression for the soft photon energy:

$$\varepsilon = \frac{m_e^2}{qE} \tag{1.37}$$

$$\Rightarrow \left(\frac{\varepsilon}{\text{eV}}\right) = \frac{0.26}{q\left(\frac{E}{\text{TeV}}\right)} \tag{1.38}$$

$f(q)$ reaches its maximum value at $q \approx 0.51$ [Vas99b]. According to equation 1.38 this results in an EBL photon energy of $\varepsilon \approx 0.5\,\text{eV}$ for a 1 TeV γ in a head-on collision.

When averaging over collision angles (detailed explanations will follow), the cross section, now described by a function $F(q)$, peaks at higher EBL energies, namely with $q \approx 0.28$ at $\varepsilon \approx 0.93\,\text{eV}$ for a 1 TeV γ-ray. Figure 1.11 shows the cross section for head-on collisions (dashed curves) and averaged over all collision angles (solid curves), without EBL absorption.

1.5.3 Absorption by the EBL

The optical depth τ describes the attenuation of the high energy (intrinsic) gamma-ray flux F'_γ.
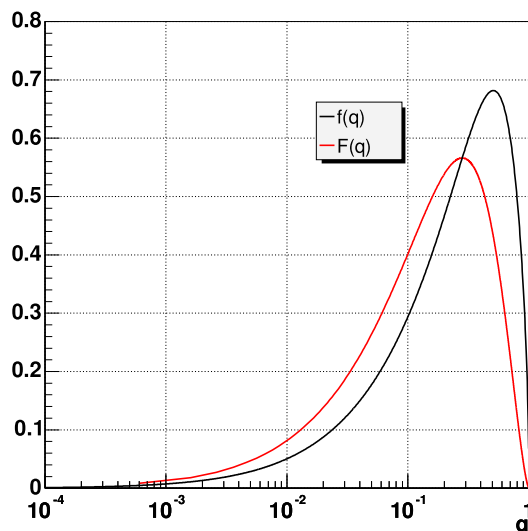
Figure 1.10: Graphs for the functions $f(q)$ and $F(q)$, which describe the energy dependence of the $\gamma\gamma$-pair production cross section (for more explanations see text, for the definition of $F(q)$ see section 1.5.3). The peak values are ~ 0.51 and ~ 0.28 for $f(q)$ and $F(q)$ respectively.

$$F_\gamma = F'_\gamma \cdot e^{-\tau} \tag{1.39}$$

with

$$\frac{d\tau}{dl} = \mu = n\sigma \tag{1.40}$$

$$\frac{dF_\gamma}{F_\gamma} = -\mu dl = -d\tau \tag{1.41}$$

μ is the familiar absorption coefficient, which depends on the cross section σ as $\mu = n \cdot \sigma$ with n being the target number density.

With $dn(\varepsilon)/d\varepsilon$ being the number of photons per unit volume per unit of energy interval, the fraction f_θ of photons moving in a differential cone at an angle θ within the interval $[\theta, \theta + d\theta]$, assuming an isotropic distribution of the EBL photons, reads:

$$f_\theta = \frac{d\Omega}{4\pi} = \frac{2\pi \sin\theta d\theta}{4\pi} = \frac{1}{2}\sin\theta d\theta = dx \tag{1.42}$$

Thus, it follows

$$n = f_\theta \frac{dn(\varepsilon)}{d\varepsilon} d\varepsilon \tag{1.43}$$

$$= \frac{dn(\varepsilon)}{d\varepsilon} d\varepsilon \, dx \tag{1.44}$$

From equation 1.40, one finds that $d\tau = n\sigma dl$. Yet, one has to keep in mind, that the EBL photons are a 'moving target'. When the VHE photon travels a distance $dl = cdt$, the EBL photons move the distance $cdt\cos\theta$ (taking the VHE photon's track as reference axis), the difference being

$$dl' = (1 - \cos\theta) \cdot dl = 2x\, dl \tag{1.45}$$

Therefore, one obtains [GS67]

$$d\tau = \sigma \frac{dn(\varepsilon)}{d\varepsilon} d\varepsilon 2x dl dx \tag{1.46}$$

Using equation 1.33 for $\sigma_{\gamma\gamma}$, one finds (see also [Ste99] and [Vas99b]):

$$\frac{d\tau}{dl} = \frac{3}{8}\sigma_T \int_{\frac{m_e^2}{E}}^{\infty} d\varepsilon \int_{\frac{m_e^2}{E\varepsilon}}^{1} 2x dx \frac{dn(\varepsilon)}{d\varepsilon} f\left(\frac{m_e^2}{E\varepsilon x}\right) \tag{1.47}$$

Assuming an isotropic EBL, an integration over the collision angles yields [Vas99b]:

$$\frac{d\tau}{dl} = \frac{3}{8}\sigma_T \int_{\frac{m_e^2}{E}}^{\infty} d\varepsilon \frac{dn(\varepsilon)}{d\varepsilon} F(\tilde{q}) \tag{1.48}$$

$$F(\tilde{q}) = 2\tilde{q}^2 \int_{\tilde{q}}^{1} f(q)\, q^{-3}\, dq \tag{1.49}$$

$$\tilde{q} = \left(\frac{m_e^2}{E\varepsilon}\right) \tag{1.50}$$

$$\tag{1.51}$$

For cosmological distances, the following redshift dependencies must be taken into account (variables referring to redshift $z = 0$ have a subscript 0).

- *Energies of EBL and VHE photons at redhift z*

$$E(z) = (1+z)E \tag{1.52}$$
$$\varepsilon(z) = (1+z)\varepsilon \tag{1.53}$$

  The integration over the soft photon energy starts at the threshold energy (which is related to z=0)

$$\varepsilon_{0,thresh} = \frac{m_e^2}{E(1+z)^2} \tag{1.54}$$

- *Density at redshift z [Ste69, Ste72]* The radius of the universe at redshift z can be expressed as:

$$\frac{R(z)}{R_0} = \frac{1}{1+z} \tag{1.55}$$

  Thus, it follows

$$n(z) = (1+z)^3 n_0 \tag{1.56}$$

- *Cosmological redshift-distance relation [Ste69]*

$$\frac{dl}{dz} = \frac{c}{H_0} \frac{1}{(1+z)^2(1+\Omega z)^{1/2}} \tag{1.57}$$

Here H_0 is the Hubble parameter and $\Omega = \frac{\rho}{\rho_{crit}}$ is the ratio of the mean density of matter in the universe to the critical density.

The correct expression for the optical thickness τ including the redshift dependencies listed above reads [Vas99b] (see also [SdJ95]):

$$\tau(E,z) = \frac{3}{8} \sigma_T \frac{c}{H_0} \int_0^z \sqrt{1+z'}\, dz' \int_{\frac{m_e^2}{E(1+z')^2}}^{\infty} d\varepsilon \frac{dn(\varepsilon)}{d\varepsilon} F\left(\frac{m_e^2}{E\varepsilon(1+z')^2}\right) \tag{1.58}$$

The integration over z extends from 0 to the maximum redshift, which is the redshift z of the source. The equation given above is derived under the assumption that the cosmological density parameter Ω equals 1. Furthermore, evolutionary effects of the EBL are neglected, which is justified for redshifts $z < 0.3$ [Vas99b].

However, the knowledge about the cosmological redshift-distance relation and the EBL density have changed during the recent scientific research:

- In the recent years the so-called *cosmic concordance model* has been established, which favors the values $\Omega_\Lambda \approx 0.7$, $\Omega_M \approx 0.3$ and $\Omega_K = 0$ (vanishing spatial curvature, flat universe). For further information on the above given cosmological density parameters see e.g. [R+04] and [Bla04].

- Assuming a conserved number of photons the present day spectral density $n(\varepsilon)$ can be calculated as given in equation 1.56 in order to obtain the spectral density at redshifts $z > 0$. This consideration is correct if the EBL has been produced mostly at higher redshifts, for example due to an initial short burst of star formation. However, such a scenario can be ruled out with the presently available EBL estimates derived from galaxy counts [Bla04].

Taking into account the just mentioned facts one must introduce following generalized redshift-distance relation [BM04] and replace equation 1.57:

$$\frac{dl}{dz} = \frac{c}{H_0} \frac{1/(1+z)}{[\Omega_M(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda]^{\frac{1}{2}}} \tag{1.59}$$

Using the parameters of the concordance model given above yields

$$\frac{dl}{dz} = \frac{c}{H_0} \frac{1/(1+z)}{[0.3 \cdot (1+z)^3 + 0.7]^{\frac{1}{2}}} \tag{1.60}$$

Furthermore, in order to account for evolutionary effects of the EBL, an interpolative function $n(\varepsilon, z)$ was introduced making use of the EBL models provided by [KBMH04] for each successive step $\Delta z = 0.1$. Figure 1.12 shows the effect of the EBL absorption for a source at $z = 0.3$ calculated in this way ('model C') in comparison with other models.
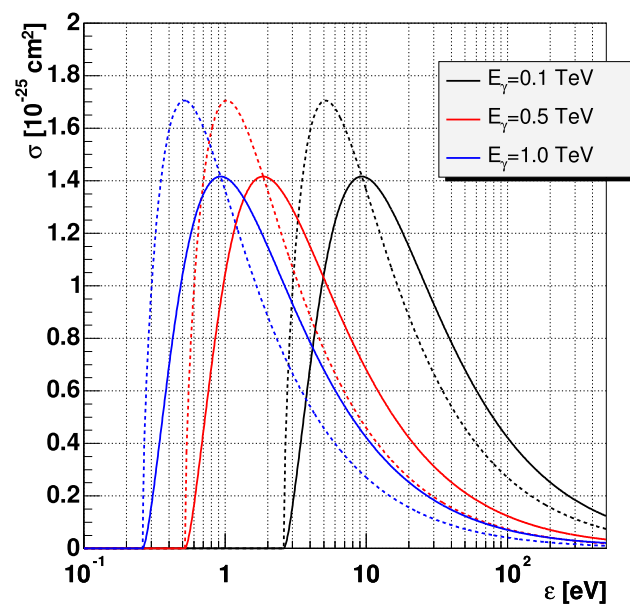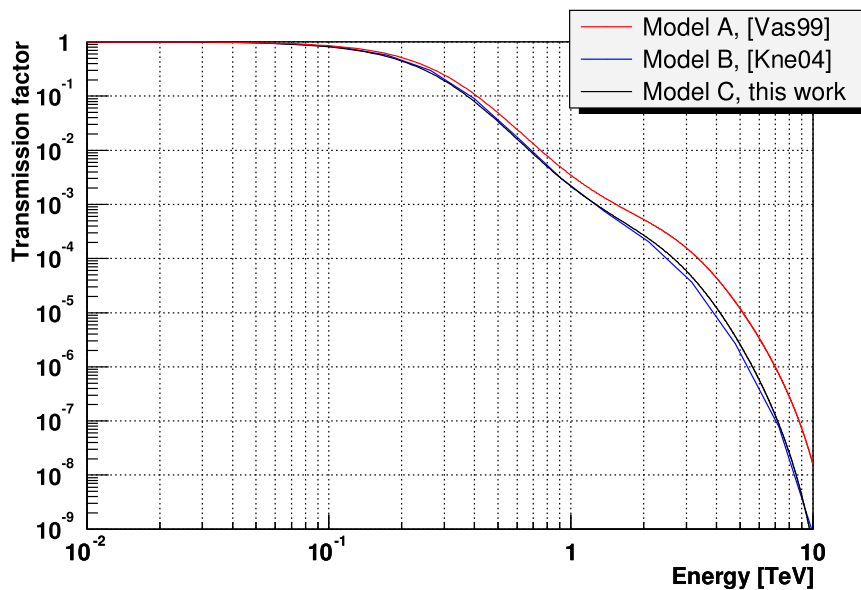
Figure 1.11: Cross sections calculated according to equation 1.33 as function of the soft photon energy ε for different gamma energies. The dashed curves are for head-on collisions whereas the solid curves describe the cross sections averaged over the collision angle assuming an isotropic distribution of the soft photons.

There is a good agreement between model C and model B whereby model B denotes the direct usage of tabulated values of the optical thickness τ as given in [KBMH04]. Model A completely follows [Vas99b], i.e. equation 1.58, using the present day EBL density given in [KBMH04].

It shall be mentioned that for nearby sources like Mkn421 and Mkn501 ($z \approx 0.03$) the redshift-distance relation can be simplifiedly expressed as:

$$l = \frac{\beta c}{H_0} \approx \frac{zc}{H_0} \ , \ z \ll 1 \tag{1.61}$$

This equation may be useful for approximate calculations.

Figure 1.12: Transmission factor $e^{-\tau}$ as a function of the VHE photon energy for a source with redshift $z = 0.3$. The absorption already shows strong effects at photon energies $E_\gamma \gtrsim 200$ GeV. Model A denotes the optical thickness calculated according to equation 1.58. For model B the optical thickness was directly taken from tables provided in [KBMH04], and model C corresponds to the method discussed in this section taking into account the concordance model and evolutionary effects of the EBL.

Chapter 2

# Air showers and Imaging Air Cherenkov Telescopes (IACTs)

Hadrons and gammas impinging on the earth's atmosphere induce showers with different characteristics. A primary gamma-ray will interact with a nucleus of the atmosphere by means of pair-production (a photo-nuclear reaction is also possible but less probable, the ratio of the cross sections is about $\sigma_{photo}/\sigma_{pair} \approx 1/300$). The mean radiation length of gamma-rays in air is about $x_0 = 36 \frac{\mathrm{g}}{\mathrm{cm}^2}$. The charged lepton pair (an electron and a positron, muonic pair production is suppressed due to the 200 times larger muon mass) radiates bremsstrahlung photons with a typical spectrum $\frac{dN_{photon}}{dE_{photon}} \sim \frac{1.}{E_{photon}}$, which still gives rise to a large number of hard photons. These hard photons again undergo pair production processes and as a result, an electromagnetic cascade or shower develops. The number of particles increases with each interaction step until the energy of the electrons and positrons decreases below the critical energy E_c (ca. 83 MeV for electrons and positrons in air), where the energy loss due to excitation and ionization dominates over the bremsstrahlung process. All the secondary particles are produced with tracks having angles well below 1° with respect to the primary particle track (the mean scattering angle is around 0.5° [Gai90]). Electrons and positrons travelling at a velocity exceeding the velocity of light in the atmosphere radiate the so-called Cherenkov light with the spectral distribution $\frac{dN}{d\lambda} \sim \frac{1}{\lambda^2}$ (more details concerning Cherenkov radiation will be given in section 2.1).

On the other hand a primary hadron undergoes an interaction with the gas of the atmosphere producing a variety of secondary particles like protons, neutrons, pions, kaons and muons, electrons, photons, neutrinos as well as the corresponding anti-particles. The hadronic interaction length in air is $\lambda_h = 90 \frac{\mathrm{g}}{\mathrm{cm}^2}$. This is somewhat larger compared to the gamma-ray radiation length and as a result the first interaction point is much more fluctuating for primary hadrons than for gammas of the same energy. A shower develops because the secondary particles lead to additional interactions and decay products, which in turn may interact with the nuclei of the air or decay.

The secondary particles produced in a hadronic shower can be divided into a hadronic $(\pi^0, \pi^{\pm}, K^0, K^{\pm})$, a muonic $(\mu^{\pm})$ and an electromagnetic component $(e^{\pm}, \gamma)$, which is mostly fed by decaying π^0. The hadronic part has a comparably big lateral spread, since in hadronic interactions large transverse momenta in the order of $0.3 \mathrm{GeV}/c$ are transferred

to the produced secondary particles (see figure 2.1).

Since in hadronic showers a large fraction of the energy is shared by hadrons and neutrinos, which produce no or a little Cherenkov light, the 'Cherenkov yield' is significantly lower than for pure electromagnetic showers [Ong98]. For primary energies between 100 GeV and 10 TeV the ratio ρ of the Cherenkov light intensity found in a gamma-ray shower to that found in a proton shower of the same energy is between two and three [Feg97]. Below energies of 100 GeV a degradation sets in and the ratio ρ increases dramatically (see figure 2.2). This degradation is due to a large fraction of the shower particles being produced below the Cherenkov threshold [Ong98].
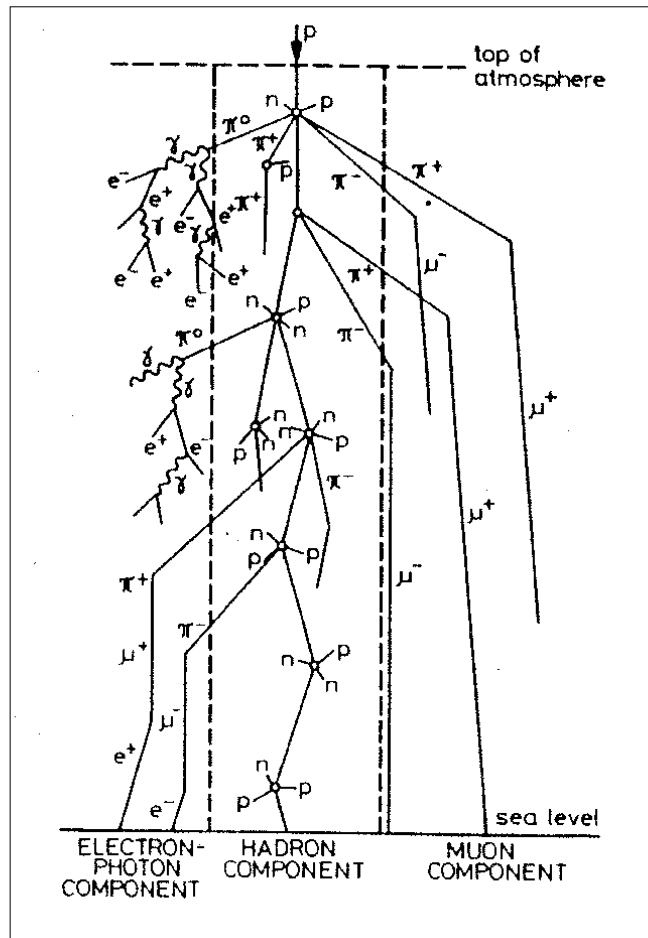
Figure 2.1: Scheme of an air-shower as produced by a primary hadron. The shower can be divided into a hadron, a muon and an electron-photon component. Taken from [AG84].

The lateral spread of hadronic showers is affecting the whole air shower structure, leading also to electromagnetic subshowers at some distance from the primary particle's track. The larger lateral spread and more fragmented structure of a hadronic shower in comparison with a pure electromagnetic shower (as it is produced by a primary gamma-
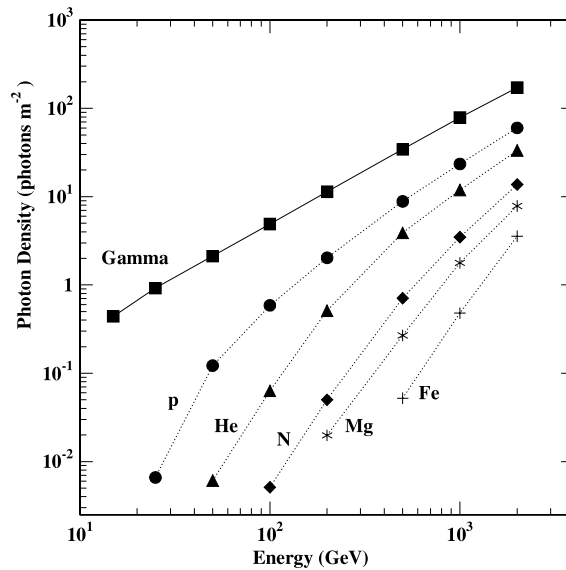
Figure 2.2: Cherenkov photon density at an observation level of 2000 m a.s.l. for different primaries. The photon density is averaged over an area of $50000 \, \text{m}^2$. Taken from [Ong98].

ray) is reflected in the Cherenkov light distribution (see figure 2.3), and therefore also in a Cherenkov light image observed by an imaging air Cherenkov telescope (see figure 2.4).

Thus, the main idea in using *imaging* Cherenkov telescopes is, that differences in the development of hadronic and gamma showers lead to the characteristic differences in the corresponding Cherenkov light images, and one is able to reject the overwhelming background of hadrons by means of an image shape analysis.

2.1 Air shower development and emission of Cherenkov light

In the following, electromagnetic showers are described in more detail by means of semi-analytical approaches.

Alternating pair production and bremsstrahlung processes in electromagnetic showers lead to a particle multiplication at each step. In a simple toy model [Gai90] one can think of a duplication of the number of particles after one radiation length. The shower maximum is reached, when the particles' energies fall below the critical energy. Then ionization and excitation processes will dissipate the remaining energy, the shower dies out.

In the case of electromagnetic cascades, there are simple approximate solutions of the cascade equations, which describe the shower development. For this purpose it is useful to introduce the so-called shower age.
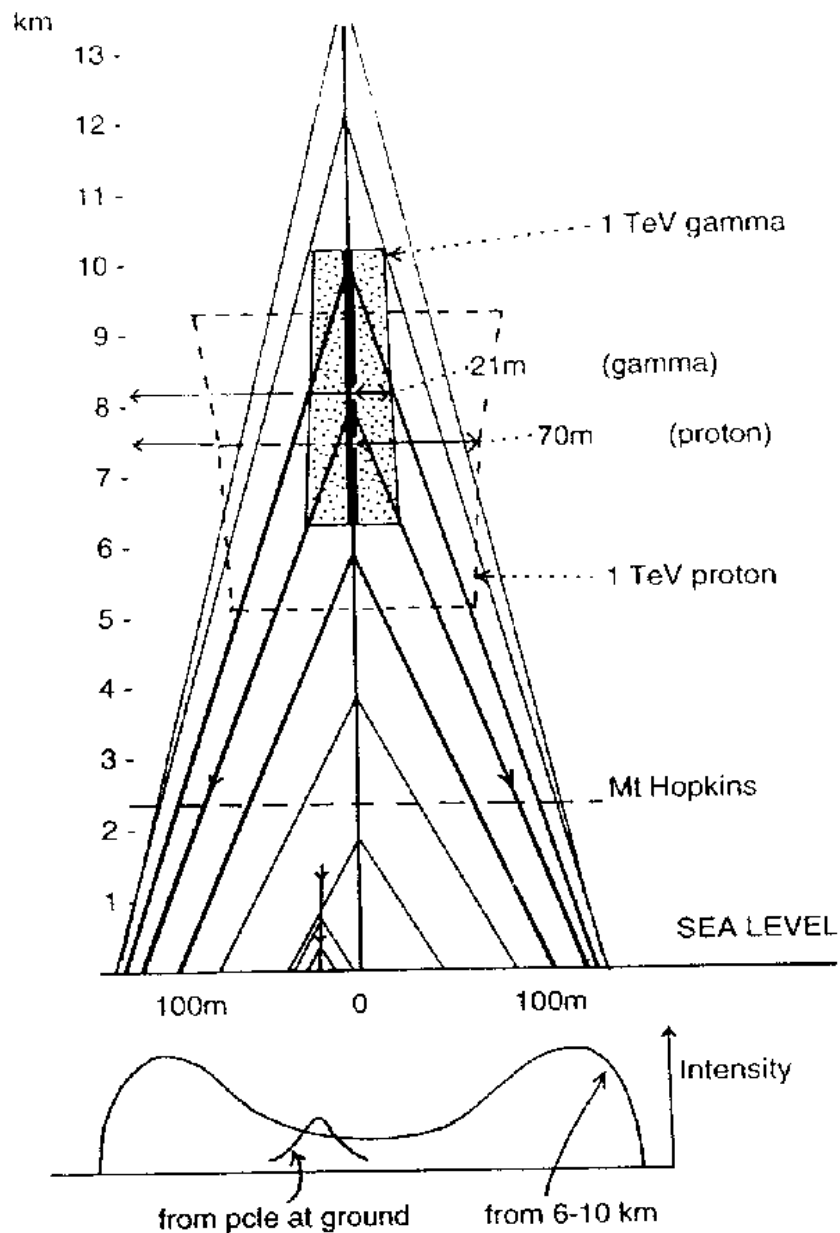
Figure 2.3: Scheme showing the geometry and shape of the Cherenkov light pool. The hump in the Cherenkov light pool at ground level corresponds to Cherenkov photon emission at altitudes of $8 - 10$ km. The solid and dashed frames indicate the areas of maximum Cherenkov light emission for gamma and proton initiated showers respectively. Taken from [Hil96].
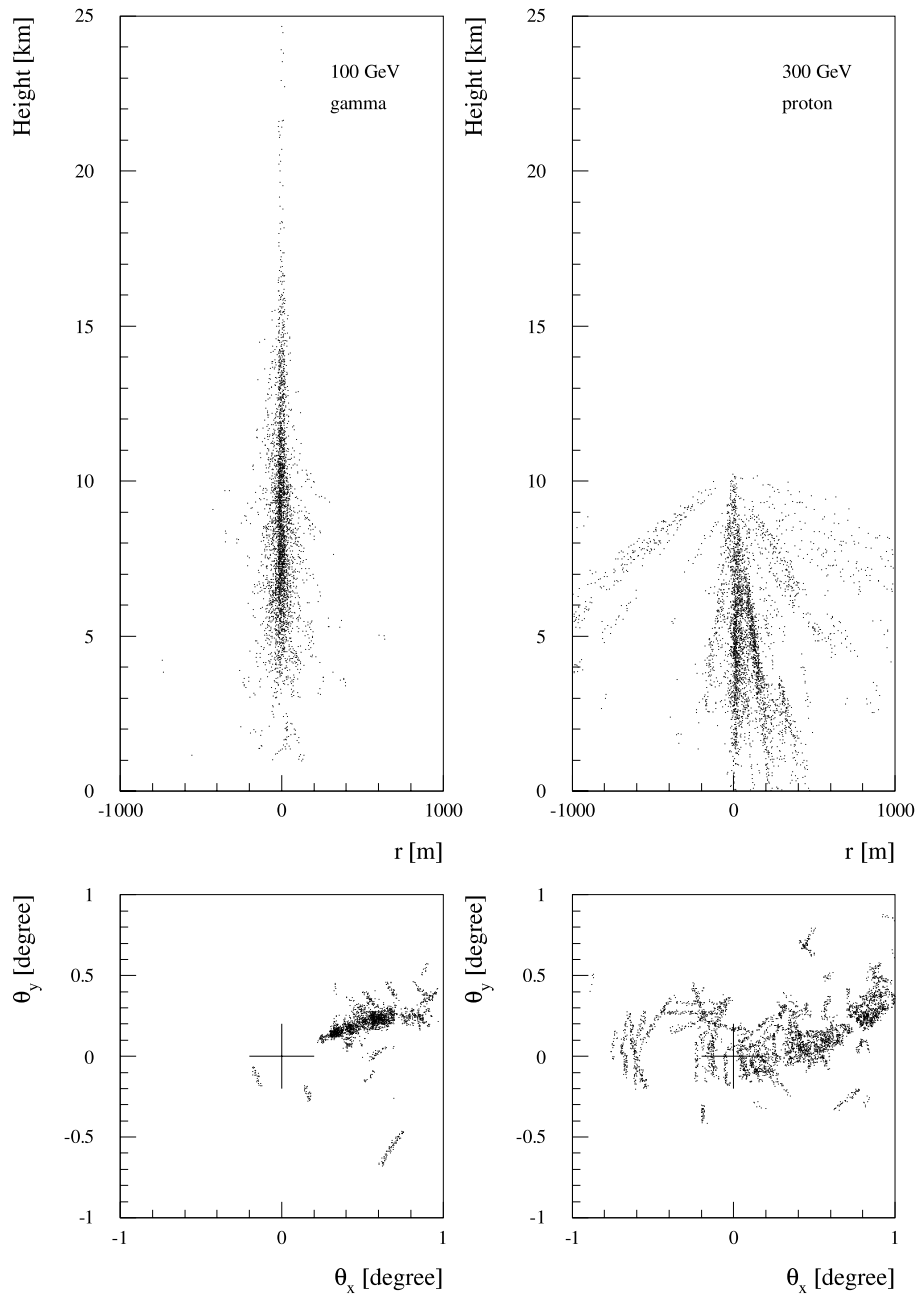
Figure 2.4: Monte Carlo simulation of a gamma and a proton induced cascade. The proton shower was chosen with an energy 3-times larger than the gamma energy, so that the Cerenkov photon output is comparable. Image taken from [A$^+$97].

$$s = \frac{3}{1 + 2 \cdot \frac{y}{T}} \qquad (2.1)$$

Here y is given by

$$y = \ln(\frac{E_\gamma}{E_c}) \qquad (2.2)$$

with E_γ being the primary gamma's energy and $E_c = 83$ MeV the critical energy in air. Another parameter T is introduced, which is the shower depth in units of radiation lengths λ. Assuming an isothermal atmosphere, T can be written as:

$$T = T_0 \cdot e^{-\frac{H}{H_0}} \qquad (2.3)$$

T_0 is calculated as $\frac{X_0}{\lambda cos\theta}$ with $X_0 \approx 1013\,\text{g/cm}^2$ being the column height of air (also called mass overburden) at sea level, θ being the zenith angle and $H_0 \approx 8400$ m being the scale height. The exponential growth in the number of particles as function of the altitude (longitudinal shower development) is given by [Gai90]:

$$N_e(s) = \frac{0.31}{\sqrt{y}} e^{T(1 - 1.5 \cdot ln(s))} \qquad (2.4)$$

There is also an analytical description of the lateral shower distribution, which holds for $0.5 \leq s \leq 1.5$ and is known as the Nishimura-Kamata-Greisen-, or shortly NKG-formula [Gre60]:

$$f(r) = \frac{\Gamma(4.5 - s)}{2\pi\Gamma(s)\Gamma(4.5 - 2s)} \left(\frac{r}{r_M}\right)^{s-2} \left(1 + \frac{r}{r_M}\right)^{s-4.5} \qquad (2.5)$$

As can be seen, the lateral distance r is scaled with the Moliere radius r_M, where r_M is given by

$$r_M \approx 9.3\frac{\text{g}}{\text{cm}^2} \qquad (2.6)$$

For sea level altitude one obtains a Moliere radius r_M of approximately 78 m.
Now, one can write down the shower-density of electrons as a function of T and lateral distance r as follows:

$$\rho_N(r, T) = \frac{N_e(T)}{r_M^2} f(r) \qquad (2.7)$$

The fraction of charged particles (i.e. electrons and positrons) traveling at 'superluminal' speed, is emitting a characteristic light - the already mentioned Cherenkov light. Even though the charged particles of the extensive air shower may not reach ground, the Cherenkov light can be utilized for reconstructing the air shower.

The 'Cherenkov condition' (charged particles moving at superluminal speed) can be met in all media with a refractive index $n > 1$, in particular 'natural' media like ice (e.g. neutrino telescopes), water (neutrino telescopes, water Cherenkov detectors for air shower investigation) and air (atmospheric Cherenkov telescopes). The diagrams in figure 2.5 illustrate the origin of the Cherenkov light. The charged particle - moving at constant

speed - causes an instantaneous polarization of the transmissive medium (dipoles are 'created' instantaneously, c.f. figure 2.5, see also [Gru93, Gru96]). As long as the particle moves at a low speed $v < c/n$ the net field is zero. If the particle's velocity exceeds the light speed c/n, the arrangement of the dipoles becomes asymmetric as depicted in figure 2.5 and a resulting field is created.
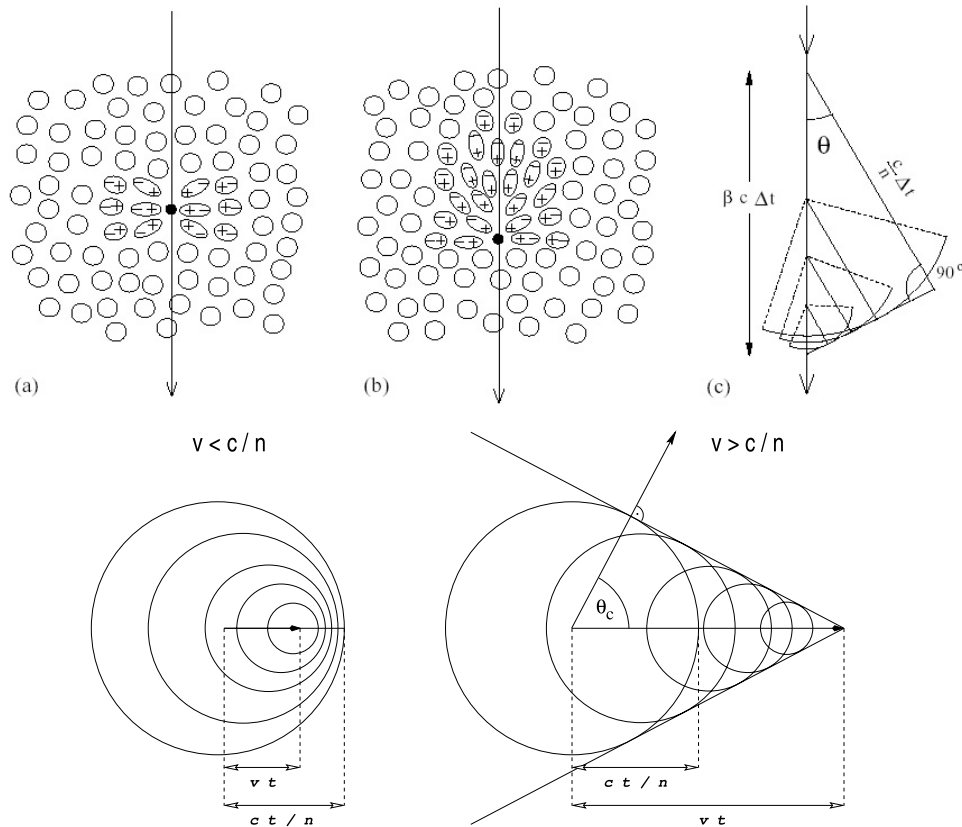
Figure 2.5: Top: The emission of Cherenkov light can be explained by investigating the polarization of the surrounding medium. If the charged particle travels at a speed below the speed of light in the medium a symmetric polarization occurs, and thus the dipole fields add up destructively (a). If the particle moves at superluminal speed the polarization becomes asymmetric and a resulting electromagnetic field forms (b). **Bottom:** Huygens' construction for determining the direction, in which the wavefront of the Cherenkov radiation propagates. In this way, Cherenkov light is interpreted as 'shock wave' being produced by the particle moving at superluminal motion - in analogy to the well-known phenomenon of a supersonic shock front. A cone is formed, since only at this geometrical position the wavefronts interfere constructive.

Huygens principle can be utilized in order to derive the characteristic angle of the Cherenkov cone:

$$\cos\theta = \frac{1}{n\beta} \qquad (2.8)$$

Here, n denotes the refractive index of the traversed medium and β is the particle's velocity in units of vacuum light speed. A more detailed derivation taking into account the recoil due to the emitted Cherenkov photon (a simple kinematic calculation) yields:

$$\cos\theta = \frac{1}{n\beta} + \frac{\hbar k}{2\,p}\left(1 - \frac{1}{n^2}\right) \tag{2.9}$$

The second term can be neglected since the photon's momentum $\hbar k$ is much smaller than the particle's momentum p.

A theoretic explanation of the Cherenkov effect (more precisely called Vavilov-Cherenkov effect after its co-discoverer) was first given by Frank and Tamm (see e.g. [Eva55]). They derived the number of Cherenkov photons per track length dx:

$$\frac{dN}{dx} = 2\pi\,\alpha\,z^2 \int_{\lambda_1}^{\lambda_2}\left(1 - \frac{1}{n^2\beta^2}\right)\frac{1}{\lambda^2}\,d\lambda \tag{2.10}$$

$$= 2\pi\,\alpha\,z^2 \int_{\lambda_1}^{\lambda_2} \sin^2\theta\,\frac{1}{\lambda^2}\,d\lambda \tag{2.11}$$
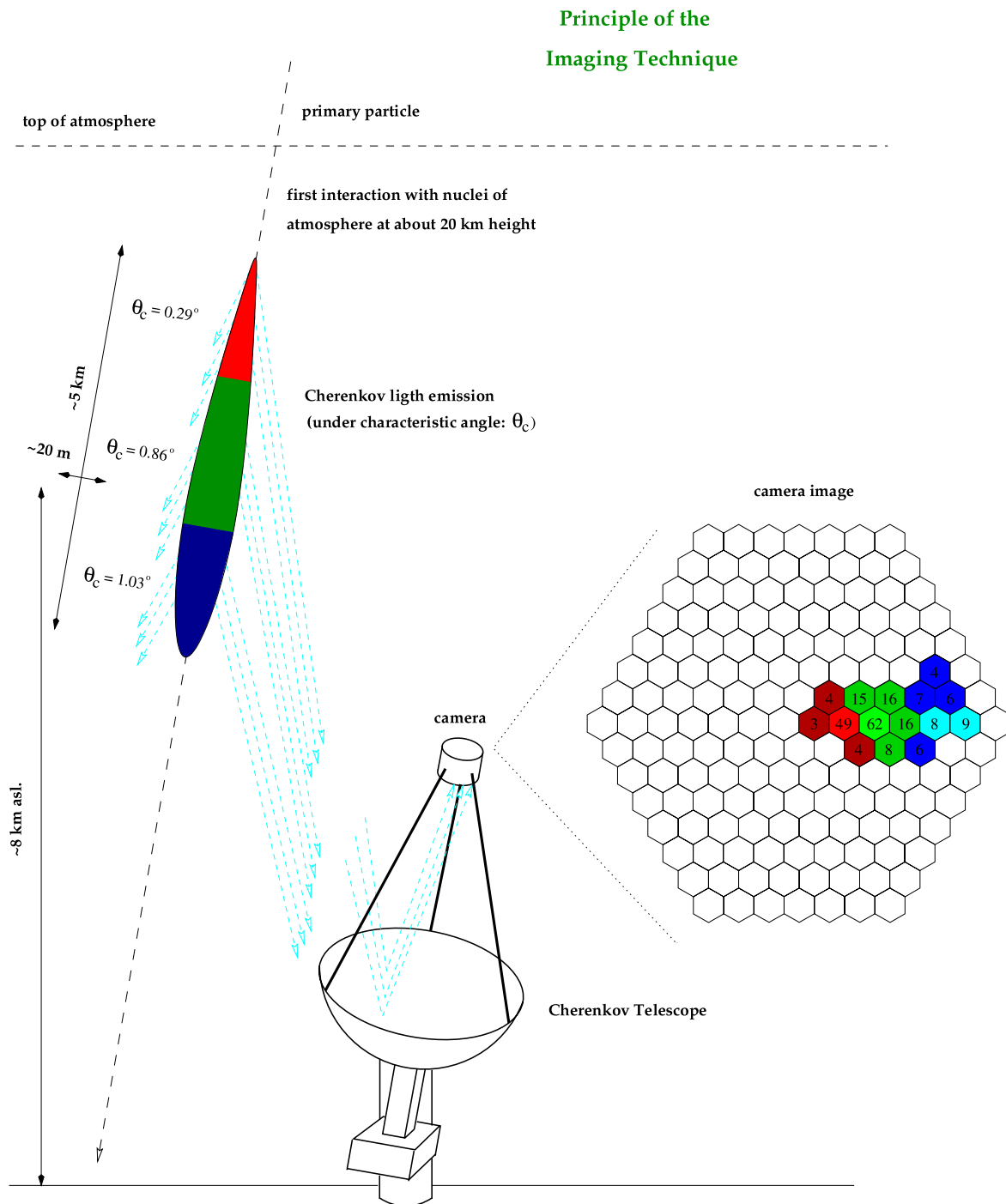
α is the finestructure constant, z the particle's charge and λ the wavelength of the emitted Cherenkov photons. Neglecting dispersion yields the equation below, which can be used to derive the number of Cherenkov photons (inside the detector's spectral sensitivity range) per track length.

$$\frac{dN}{dx} = 2\pi\,\alpha\,z^2\sin^2\theta\frac{\lambda_2 - \lambda_1}{\lambda_2\,\lambda_1} \tag{2.12}$$

2.2   The Imaging Technique

The shower images in the cameras of Cherenkov telescopes are commonly described by simple shape and orientation/position parameters. Therefore one introduces orientation angles, the distance between the source position (in normal observation mode identical with the camera center) and the image center of mass and the length and width of the image, which are related to the image's second moments via a principal axis transformation. See figure 2.6 for the imaging geometry and figure 2.7 for an explanation of the image parameters, the so-called Hillas Parameters [Hil85].

After a cleaning of the image (see chapter 4), which removes noise dominated pixels, the image analysis proceeds in calculating the image's moments. In the following formulae, the sums run over all pixels in the camera. The weight w_i for pixel i is taken to be the number of Cherenkov photons q_i recorded by this pixel divided by the total number of recorded Cherenkov photons .

Principle of the

Imaging Technique

top of atmosphere

primary particle

first interaction with nuclei of
atmosphere at about 20 km height

$\theta_c = 0.29°$

~5 km

~20 m

$\theta_c = 0.86°$

$\theta_c = 1.03°$

Cherenkov ligth emission
(under characteristic angle: θ_c)

camera image

~8 km asl.

camera

	4	15	16	7	6
3	49	62	16	8	9
	4	8	6		

4

Cherenkov Telescope

Figure 2.6: This drawing illustrates the formation of the Cherenkov light image in the
camera plane. Cherenkov photons emitted at an early stage in the shower development
(red marked region) are 'seen' under a smaller angle than photons emitted at the shower
maximum (green marked) or shower tail (blue marked). Also shown is the (charge)
asymmetry of the Cherenkov light image (the numbers inside the camera pixels denote
the charge in units of photoelectrons). Taken from [Kra01].
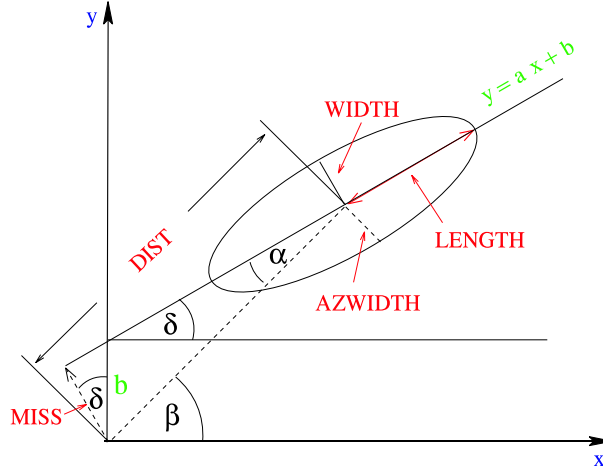
Figure 2.7: Scheme showing the Hillas parametrization [Hil85] of the Cherenkov light image. The basic idea is, that the recorded 2-dimensional charge distribution is described by an ellipse. The axes of this Hillas ellipse are the second moments of the charge distribution in the coordinate system, where the correlation vanishes (the covariance matrix is diagonalized, see text for explanations). The Hillas Parameters *width*, *length*, *dist* and *alpha* (depicted as α) are indicated. Also the additional parameters *azwidth* and *miss* are shown.

$$w_i = \frac{q_i}{\sum_j q_j} \tag{2.13}$$

$$<x> = \sum_i w_i \cdot x_i \tag{2.14}$$

$$<x^2> = \sum_i w_i \cdot x_i^2 \tag{2.15}$$

$$<y> = \sum_i w_i \cdot y_i \tag{2.16}$$

$$<y^2> = \sum_i w_i \cdot y_i^2 \tag{2.17}$$

$$<xy> = \sum_i w_i \cdot x_i \cdot y_i \tag{2.18}$$

With the definitions given above, one can calculate the second moments and the covariance (here denoted as σ_{xy}).

$$\sigma_x^2 = <x^2> - <x>^2 \tag{2.19}$$
$$\sigma_y^2 = <y^2> - <y>^2 \tag{2.20}$$
$$\sigma_{xy} = <xy> - <x><y> \tag{2.21}$$

Thus, one yields the covariance matrix C.

$$C = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \tag{2.22}$$

In order to disentangle orientation and shape parameters (the second moments are describing the image 'spread' and are thus referred to as shape parameters) a principal axis transformation using a rotation matrix M is performed:

$$C' = M C M^T \tag{2.23}$$

$$M = \begin{pmatrix} \cos\delta & \sin\delta \\ -\sin\delta & \cos\delta \end{pmatrix} \tag{2.24}$$

The covariance matrix C' in the rotated coordinate system (related variables are primed in the following) must take on diagonal form. The ansatz $M C M^T = \begin{pmatrix} \sigma_{x'}^2 & 0 \\ 0 & \sigma_{y'}^2 \end{pmatrix}$ leads to a quadratic equation in $\tan\delta$ with the solutions

$$\tan\delta = \frac{d \pm \sqrt{d^2 + 4\sigma_{xy}^2}}{2\sigma_{xy}} \tag{2.25}$$

$$d = \sigma_x^2 - \sigma_y^2 \tag{2.26}$$

An ambiguity in the sign of the above given equation arises, since there are two axes of the rotated image and thus two possible rotation angles differing from each other by $90°$. It is sufficient to restrict oneself to the positive sign, which corresponds to the x-axis of the rotated coordinate system being along the main axis of the image (see [Wit02a], this can be proven by deriving $\sigma_{x'}^2 - \sigma_{y'}^2 > 0$ using the equations for *width* and *length* given below). One can furthermore restrict the domain of δ to $[-90°, 90°]$. The Hillas Parameters *length* and *width* are now given as follows:

$$length = \sigma_{x'} \tag{2.27}$$

$$width = \sigma_{y'} \tag{2.28}$$

This can be expressed using only (unprimed) quantities of the camera coordinate system.

$$length = \sqrt{\cos^2\delta\,\sigma_x^2 + \sin^2\delta\,\sigma_y^2 + 2\sin\delta\,\cos\delta\,\sigma_{xy}} \tag{2.29}$$

$$width = \sqrt{\cos^2\delta\,\sigma_y^2 + \sin^2\delta\,\sigma_x^2 - 2\sin\delta\,\cos\delta\,\sigma_{xy}} \tag{2.30}$$

$$dist = \sqrt{<x>^2 + <y>^2} \tag{2.31}$$

It is helpful to introduce the image axis $y = ax + b$ (see figure 2.7), whose parameters a and b can be calculated easily:

$$a = \tan\delta \tag{2.32}$$

$$b = <y> - a <x> \tag{2.33}$$

miss is the perpendicular distance from the camera center to the the image axis. One finds

$$miss = b\cos\delta \tag{2.34}$$

$$= \frac{b}{\sqrt{1+a^2}} \tag{2.35}$$

The so-called *alpha*-parameter is the angle between the image main axis and the line, which runs through the camera center and the image's center of gravity.

$$alpha = \sin^{-1}\left(\frac{miss}{dist}\right) \tag{2.36}$$

$$= \sin^{-1}\left(\frac{b}{dist \cdot \sqrt{1+a^2}}\right) \tag{2.37}$$

It is straightforward to obtain expressions for *width* and *length* eliminating δ and introducing a and b:

$$length = \sqrt{\frac{\sigma_x^2 + 2a\sigma_{xy} + a^2 \cdot \sigma_y^2}{1+a^2}} \tag{2.38}$$

$$width = \sqrt{\frac{a^2 \cdot \sigma_x^2 - 2a\sigma_{xy} + \sigma_y^2}{1+a^2}} \tag{2.39}$$

It shall be noted here, that the MAGIC camera consists of two pixel types (see next chapter). The outer pixel rings are made up by pixels 4-times bigger in area. Their recorded number of Cherenkov photons - or more general their charge - must be divided by four, which means that one deals with 'charge densities' when calculating the image parameters.

Chapter 3

The MAGIC Experiment

The MAGIC (Major Atmospheric Gamma-ray Imaging Cherenkov) telescope has been developed with the main aim to lower the energy threshold in the detection of very high energy gamma rays of existing detectors. The region right between $50\,\mathrm{GeV}$ and $300\,\mathrm{GeV}$ is beyond the upper energy limit of space borne detectors and below the energy threshold of current Cherenkov telescopes.

Satellite detectors have an upper energy limit in detecting gamma-rays, since the flux becomes too small for a reasonable data taking time, taking into account a collection area of only $\sim 1\,\mathrm{m}^2$. Furthermore, high energies make a containment within a space telescope a serious problem [Sch01]. On the other hand, imaging atmospheric Cherenkov telescopes have detection areas in the order of several $10^4\,\mathrm{m}^2$ (reflecting the area of the Cherenkov light pool), so that the upper limitation is shifted to much higher energies (saturation effects of the recording electronics and a leakage of the Cherenkov images out of the camera become dominant for $\gtrsim 10\,\mathrm{TeV}$).

The lower energy limit is due to the photon density in the Cherenkov light pool, which is proportional to the primary gamma's energy and becomes too small for low energy gammas restricting the energy range of old generation IACTs to $> 300\,\mathrm{GeV}$. See figure 3.1 for a comparison of the sensitivities in different experiments. Here, the sensitivity is defined as the minimum integral flux, which one needs in order to obtain a $5\,\sigma$ excess for a point source with a Crab-like energy spectrum ($\frac{dN}{dE} \sim E^{-2.6}$) observed under a zenith angle $0^\circ \leq \theta \leq 30^\circ$ for $50\,\mathrm{h}$.

The design of MAGIC is chosen to allow an operation at low Cherenkov photon densities by constructing a telescope with a large reflector and a highly sensitive camera. In this way one achieves:

- An energy threshold well below $100\,\mathrm{GeV}$.

- A sensitivity of $\sim 2 \cdot 10^{-11}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at $100\,\mathrm{GeV}$ (see figure 3.1)

- A high sensitivity for the VHE range in observations at large zenith angle. This is because at large zenith angles the shower maximum is further away from the telescope, which means that the Cherenkov light pool on ground covers a larger area leading to an increased effective collection area.
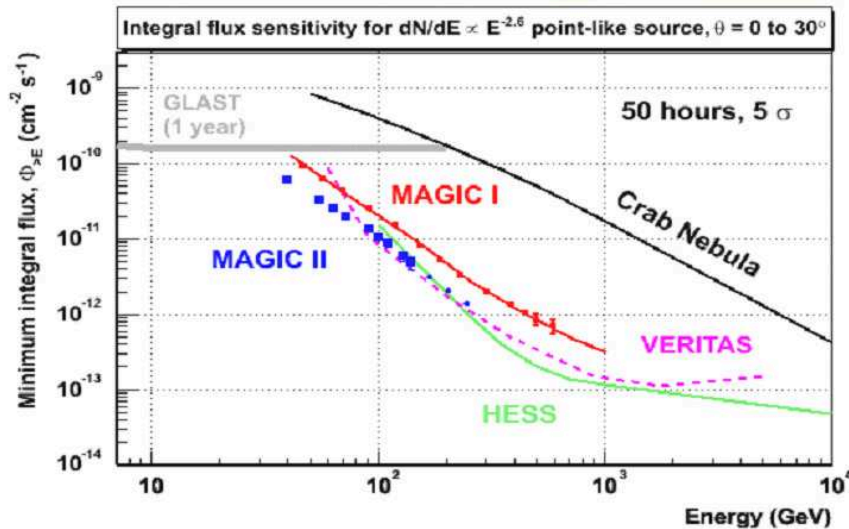
Figure 3.1: Sensitivity plot for the planned GLAST experiment (satellite detector) and for the Cherenkov telescopes Veritas, H.E.S.S. and MAGIC as well as MAGICII (shown is here the performance of the MAGIC stereo system, when the second MAGIC telescope - the so-called clone - is operational, at the time of writing this thesis the MAGIC clone was still in the construction phase). Also shown is the TeV flux from the Crab Nebula. Taken from [SM06].

3.1 The main observational aims of MAGIC

The key points for a promising operation of the MAGIC telescope are as follows:

- Observations in the up to now unexplored energy region 50 GeV − 300 GeV ('gap' region) may reveal new and unexpected spectral features.

- The detection of an increased number of sources is expected due to the higher sensitivity in comparison with previous generation IACTs.

- The higher sensitivity also leads to a better accuracy in measuring spectral and temporal features, at all accessible energies.

In the following list the main targets for MAGIC observations are compiled (see [B+97]).

- *Active Galactic Nuclei (AGNs)*
 About 3 % of all galaxies are active [Sch01]. They show a high luminous and variable, non-thermal emission from their core regions. Especially interesting is the Blazar subclass (see chapter 1). Here the main objectives are:

  – Provide a large sample of AGNs as basis for good statistical evaluations.
  – Measuring spectral and temporal features of AGNs with a so far unreached accuracy.

THIRD EGRET CATALOGUE OF GAMMA-RAY POINT SOURCES
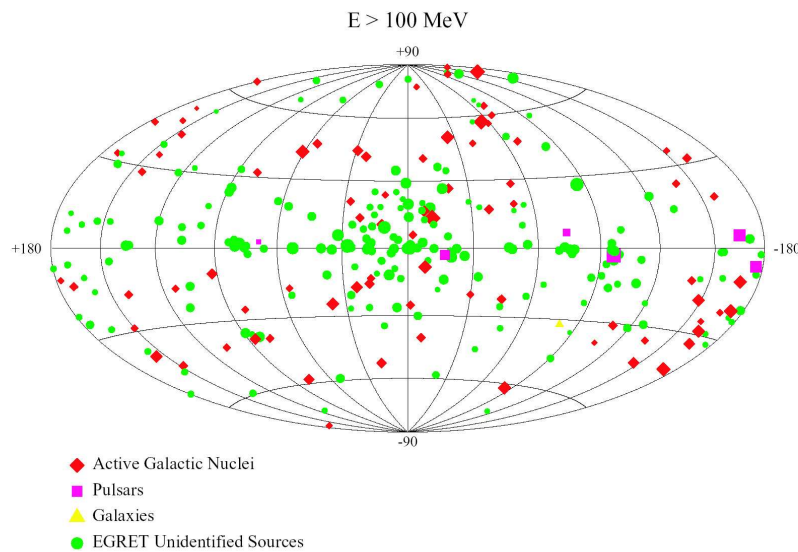
E > 100 MeV

Figure 3.2: 3rd EGRET catalogue: Sky map of γ-ray point sources in galactical coordinates for energies $100\,\mathrm{MeV} < E_\gamma < 10\,\mathrm{GeV}$. Taken from [H$^+$99].
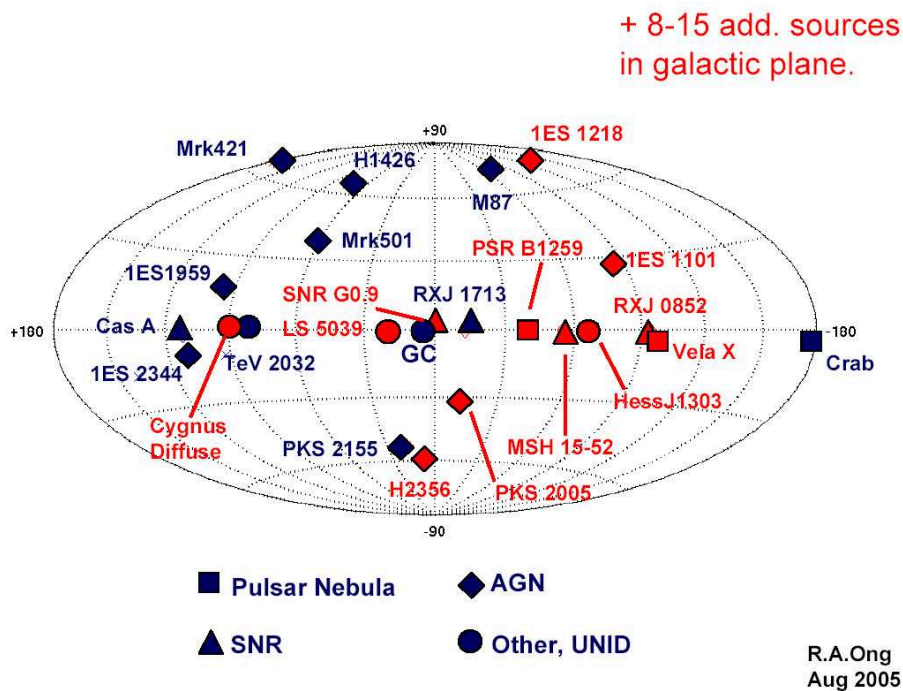
Figure 3.3: Sky map of very high energy γ-ray point sources for energies exceeding $300\,\mathrm{GeV}$. Red symbols indicate detections made after 2003. Image taken from [Ong05].

More points of particular interest are given as follows:

- Most of the AGNs detected by EGRET are expected to have an energy cut-off in the gap region, since they are not observed by the previous generation IACTs (see figures 3.2 and 3.3 showing the gamma-ray sky at energies below 10 GeV and above 300 GeV respectively). This is also suggested by the observation, that the diffuse extragalactic γ-background flux at $30\,\mathrm{MeV} - 100\,\mathrm{GeV}$ measured by EGRET has a similar spectral index as the average blazar flux. This leads to the explanation, that the diffuse γ-background is (partly) due to unresolved point sources, in particular blazars.

- A whole class of AGNs may show the second SED peak in the gap range.

- The observation of a cut-off in the gap-range can be used to constrain the density of the extragalactic photon background.

- Together with observations in other wavelength regimes one may be able to provide a better modeling of the AGN emission. Mechanisms of acceleration and hadronic/leptonic models can be studied.

- A measurement of the photon horizon (defined as the maximum cut-off energy as a function of the redshift) will provide constraints on cosmological parameters.

- *Supernova Remnants (SNRs)*
  More than 250 SNRs have been observed in our galaxy (mostly below the HE range) [Wee03], which have very different properties depending e.g. on the type of Supernova explosion and the nearby interstellar medium. They also show a wide range of ages, distances and angular sizes. Thus high sensitivity observations at all energies including the HE and VHE range will contribute to a better understanding of these objects. Observations at highest energies are particularly useful, since they usually stretch the physics models to their limits and impose strong constraints on physics parameters.

- *Stellar accretion-driven systems*
  Compact stellar objects (white dwarfs, neutron stars and black holes) are likely to produce HE and possibly also VHE γ-rays, if they are able to accrete matter. Advantageous conditions are met in binary systems, where a matter supplying non-degenerate companion star is present. Such system are:

  - *X-ray Binaries, in particular so called HMXRBs (High Mass X-ray Binaries)*, π^0-production in the region, where the matter inflow reaches the accretion disk, is very likely. Therefore a periodically modulated γ-ray signal may be observable. Multiwavelength observations including X-rays will obviously help in the interpretation of the data.

  - *Micro Quasars* are probably a subclass of X-ray Binaries, where one companion is a black hole. These objects show relativistic jets and can be seen as miniature AGNs. The development in such systems occurs much faster than in AGNs

making the observation of temporal features possible, which in AGNs take place at much larger time scales. The very strong flux variability of Micro Quasars makes high-sensitivity observations necessary.

– *Cataclysmic variables* are made up by a white dwarf and a cool M-type companion filling its Roche lobe. Here, like in the X-ray binary systems, π^0-production is likely, leading to γ-emission.

- *Pulsars*

  – *Radio Pulsars*
    Both polar cap and outer gap models have problems in describing the features of available data. Especially the measurement of the detailed structure of the uppermost end of the emission spectrum beyond the reach of the EGRET instrument is necessary to gain further insight into the pulsar acceleration mechanism.

  – *Radio-quiet Pulsars*
    So far the only known radio-quiet pulsar is the Geminga pulsar. Assuming, that it is the closest object of its type, there should be around 1600 radio-quiet pulsars in our galaxy [B$^+$97]. High sensitivity observations by MAGIC will reveal fainter Geminga-like objects not detected by EGRET. First candidates are the faint unidentified EGRET-sources having a too low signal-to-noise ratio for a periodicity analysis.

- *Gamma-ray Bursts (GRBs)* are still phenomena of basically unknown origin. The MAGIC telescope is specially designed for a fast movement and tracking of a GRB, after a satellite alert has arrived (30 s mean repositioning time, see next section). This way a direct GRB measurement is possible if the GRB-duration exceeds $10 - 100$ s. Furthermore, a possible delayed HE-component can be caught.

- *Search for Supersymmetric Particles*
  The lightest supersymmetric particle (LSP), χ, is a popular candidate for the so-called non-baryonic, cold dark matter. Its present lower mass limit is $m_\chi > 37$ GeV (95 % C.L.) as indicated by the most recent ALEPH analyses [B$^+$01]. The annihilation of such a particle with its self-conjugate may lead to gamma emission according to $\chi\overline{\chi} \to \pi^0 + \ldots \to \gamma\gamma + \ldots$ detectable by the MAGIC telescope. However there is a large uncertainty in the SUSY parameters leading to large uncertainties of both the expected continuum and line emission. Observational candidates are close-by massive objects/regions, i.e. preferentially the center of our galaxy.

- *Cosmological Magnetic Fields*
  Delayed γ's from transient phenomena (like GRBs or short duration flares of very high activity) due to cascading by $\gamma\gamma$-interactions on the diffuse extragalactic background fields and subsequent deflection of electrically charged particles in the electromagnetic cascades may serve to detect or constrain the still unknown intergalactic fields.

3.2   The location of the MAGIC telescope

The MAGIC telescope is located at the Roque de los Muchachos on La Palma, Canary Islands, Spain. The altitude is 2200 m a.s.l., the geographical coordinates are 17.89°W, 28.76°N. The Canary Islands offer one of the best observation sites in Europe. The temperature variation amounts to only $2 - 3\,^\circ$C over night. The reason is a rapid due influx of stable maritime air within one hour after sunset [B$^+$97]. Therefore absorption effects of the Cherenkov light are minimized.

The sky view of the MAGIC telescope given in galactic coordinates can be seen in figure 3.4. Regions of different zenith angle ranges are indicated by different colours. The zenith angle is calculated for the culmination point.
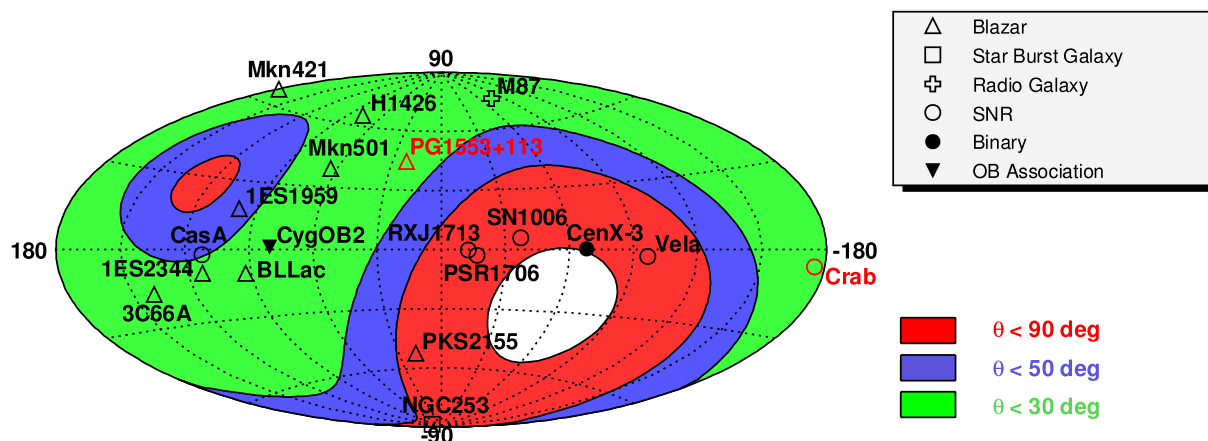

Figure 3.4: TeV gamma-ray sources as seen by the MAGIC telescope. Green, blue and red areas indicate regions of the sky observable at zenith angles below 30°, 50° and 90° respectively. Objects inside the white area are not observable since they lie below the horizon. Meaning of symbols: Triangles indicate AGNs, open circles indicate SNRs, CygOB2 is an OB association, CenX-3 is a X-ray binary system and M87 is a radio galaxy. The two red marked objects - the Crab Nebula and the BL Lac PG1553+113 - are analyzed in this work.

In the following section the technical properties of the MAGIC telescope are described in some detail.

3.3   The basic design of the telescope frame and reflector

The basic design of the MAGIC telescope was adopted from the concept of a large 17 m diameter solar concentrator, which had been built and tested already a few years ago in the frame of the German solar power research program [B$^+$97].

In figure 3.5 one can see a photo of the MAGIC telescope. The main constructional parts are indicated: The overall space frame (mainly the mirror support) is made up
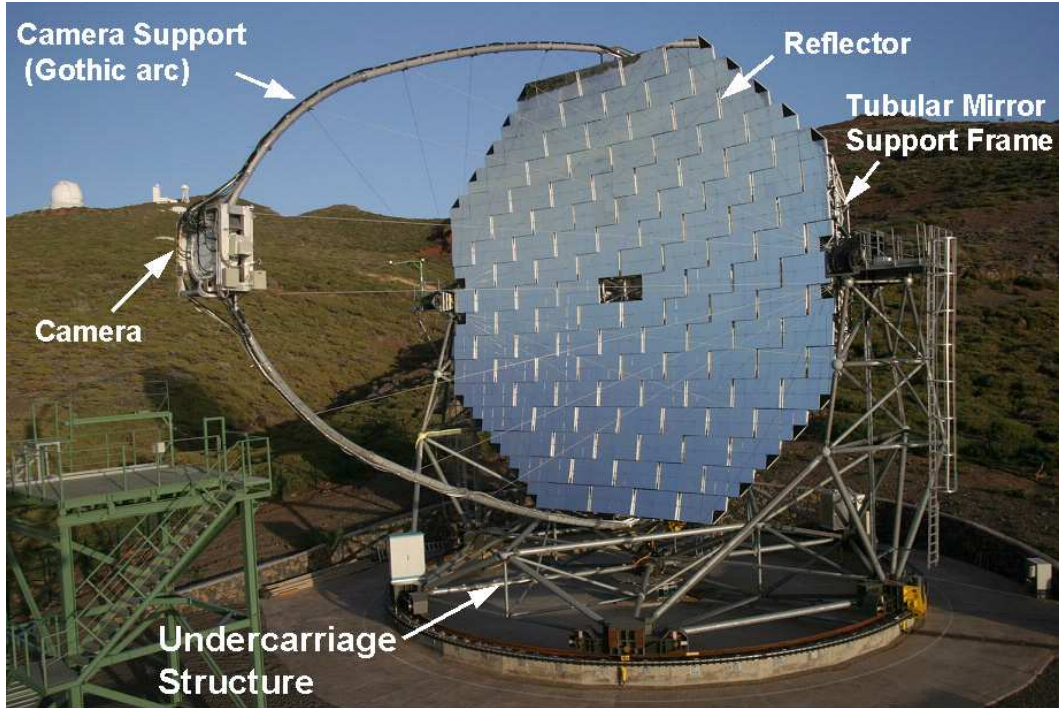
Figure 3.5: The MAGIC telescope - main constructional parts are pointed out.

by carbon fiber epoxy tubes, which combine low weight with stiffness and provide a low inertia and therefore good mobility of the telescope (in less than 30 seconds the telescope can be positioned to any skypoint [B$^+$04]). Therefore, the design of the telescope allows a fast reaction to a GRB alert by tracking the GRB sky point.

The mirror dish support is constructed as a three layer frame in a tube and knot system (no welding demanded), which results in a weight of the mirror dish plus frame of 20 tons, while the whole structure including the undercarriage (see figure 3.5) has about 60 tons [Cor04]. For the mirror dish and structure a wind resistance up to 170 km/h and a resistance against a complete ice coverage up to 3 cm are guaranteed [Pan04].

The tesselated mirror of octagonal rim-shape consists of 965 individual spherical mirrors. Groups of four (at the rim three) of these mirrors of size 49.5 cm x 49.5 cm are joined in one panel. The gross reflector dish has an approximately parabolic shape. Thus mirror elements with a bending radius of 34.125 m were chosen for a position close to the center of the mirror dish, and mirror elements with increased curvature were taken for a location at larger distances to the center up to a radius of 36.625 m for a position at the rim [BM03]. The mirror elements can be automatically adjusted to achieve a minimal point spread function (PSF). The parabolic shape preserves the intrinsic shower time spread (the individual photon arrival times inside the Cherenkov shower front are not distorted with respect to each other, when being reflected onto the camera plane).

The internal time spread and directional time evolution of the recorded shower images can be used in the analysis, in particular for gamma/hadron-separation (g/h-separation) and energy estimation. The ratio of focal length to mirror diameter f/D is close to 1

[Cor04] providing a good optical quality of the Cherenkov light images.

The aluminum mirror segments (reflecting surface) are diamond milled and quartz coated for protection reasons. They are glued on an aluminum honeycomb structure, which is held inside a aluminum box (see figure 3.6). The reflectivity has a mean value of $85\,\%$ in the wavelength range $300 - 650\,$nm and a roughness below $10\,$nm [Pan04]. In order to inhibit ice and dew deposition, a $0.3\,$mm Printed Circuit Board serving as internal heater has been build in [Cor04].
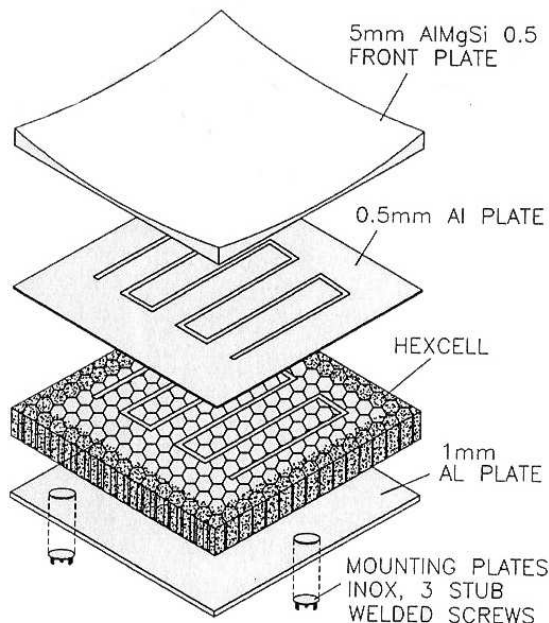

Figure 3.6: Exploded view of a mirror segment. Taken from [B$^+$97].

The large mirror with the light-weight construction makes an active mirror control (AMC) necessary, which allows a mirror adjustment during the telescope operation to compensate for small residual deformations (arising especially due to zenith angle changes). The adjustment is done by 2 stepping motors, which are additionally position controlled by means of switchable laser pointers, which produce spots on the (closed) camera lids monitored by a CCD camera (see figure 3.7).

3.4   The MAGIC Camera

The camera of a Cherenkov telescope should be particularly designed for a high sensitivity even to single (Cherenkov) photons. Therefore photomultipliers are used as sensitive photon detectors (for infrared, visible and UV range). In order to provide a complete image of a shower in its Cherenkov light, the camera consists of a photomultiplier matrix.
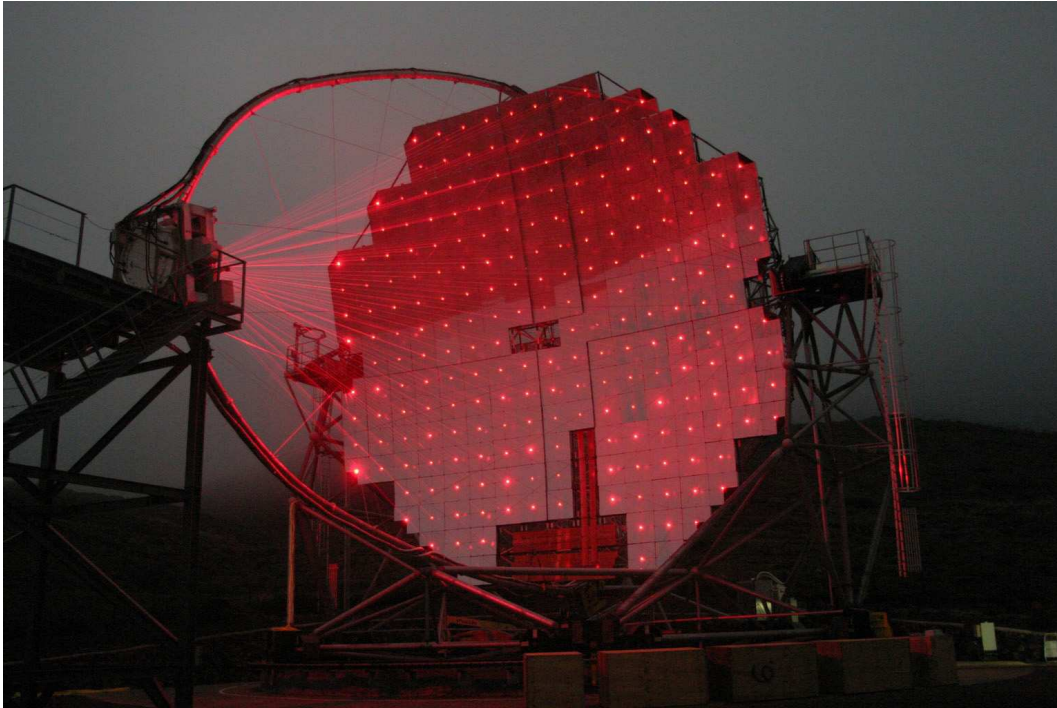
Figure 3.7: The Active Mirror Control in operation. The laser beams are visible due to scattering on mist.

This is a completely different situation than in classical astronomy where optical telescopes are equipped with high resolution but comparably low-efficiency cameras.

The MAGIC camera (as well as the reflector with its comparably big dimensions) is optimized on the detection of low-energy gamma showers. Therefore, besides the high sensitivity, which is absolutely essential, a fine pixelization of the camera (in comparison e.g. with the precursor experiment HEGRA) is chosen. This leads to:

- *Better quality image parameters*
  Besides the standard Hillas Parameters, also the charge and time asymmetry are used, whose measurement are expected to be improved by a fine pixelization. The quality of image parameter determination is an essential point in g/h-separation, when differences in the development of hadronic and electromagnetic gamma showers are exploited.

- *Lower energy threshold*
  The pixel size should be chosen according to the expected distribution of Cherenkov photons in the camera. Since at low gamma ray energies the Cherenkov light images are distributed just over a few 10th of a degree, the pixel diameter in the MAGIC camera (0.1° for the inner part, see below) is suitable.

In the following the MAGIC camera is described in more detail (see also figure 3.8): The camera consists of 577 hexagonal shaped pixels, each of which is made up of a
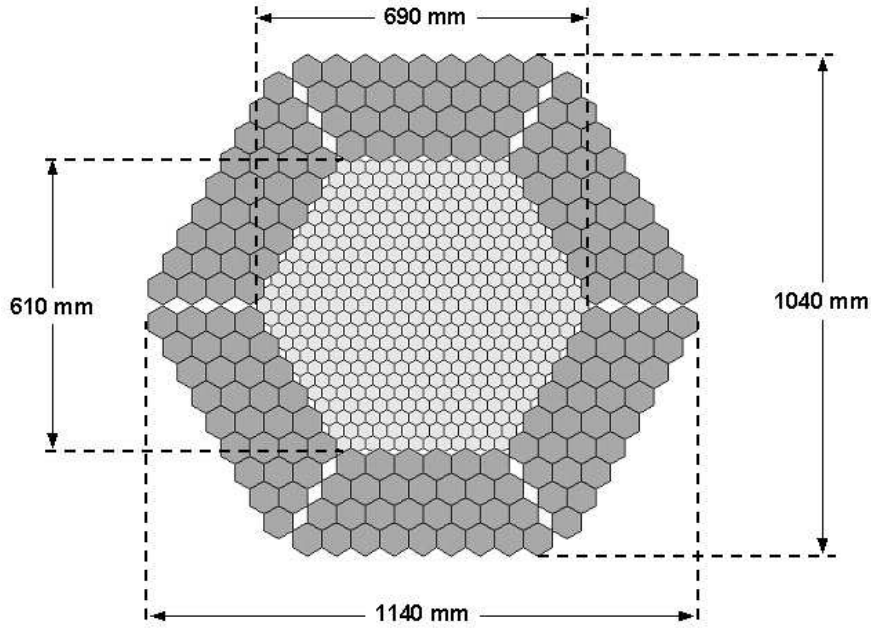
Figure 3.8: The layout of the MAGIC camera. The outer pixels are shown as dark grey filled hexagons, inner pixels are filled light grey. The camera dimensions are indicated.

photomultiplier tube (PMT) with a light funnel (the so-called Winston cone) attached to the PMT. This light funnel has a hexagonal shaped entrance to allow a close package of the pixels and ends in a circle of a diameter identical to the PMT's sensitive area. The MAGIC camera consists of a fine pixelized inner part surrounded by lower resolution pixel rings. The design was chosen mainly due to cost reasons. But there is no serious drawback, since the outer pixel rings catch mainly the shower tails and distortions due to aberration (coma) are more significant (they exceed $0.06°$ for a distance of $1°$ from the camera center [B$^+$97]) making a finer pixelization in the outer camera region not more effective. The camera setup is as follows:

- *Inner part* 397 hexagonal pixels with a diameter of 30 mm $\stackrel{\triangle}{=} 0.1°$. 325 pixels are inside the trigger zone. The field of view is $2.1° - 2.3°$.

- *Outer part* 180 hexagonal pixels with a diameter of 60 mm $\stackrel{\triangle}{=} 0.2°$. The field of view is $3.5° - 3.8°$.

The overall field of view (FOV) of $3.5° - 3.8°$ means, that an image-leakage out of the FOV is kept small up to energies of a few TeV. The Winston cones are produced from a plastic material covered with aluminized mylar foil of 85 % reflectivity. Due to the hexagonal entrance shape a coverage of nearly 100 % in contrast to 80 % (closed packed circular tubes) is achieved [Pan04]. Since the high sensitive area of the PMTs is significantly smaller than their outer dimensions even 50 % of the light would be lost [B$^+$97] without light concentrators. Moreover, due to the reduced angular acceptance of the Winston cones a reduction of large angle stray light is achieved (backscattered light from ground, e.g. moonlight or artificial light sources like flashlights or car lamps).

The plate of light concentrators in front of the PMT camera is covered by a UV-transparent plexiglas window of 2 mm thickness and high penetration, i.e. low reflectivity of 4 %. Its purpose is to hermetically seal the electronic sensors and instruments against humidity and dust. The photomultipliers of the camera are the actual detectors and so they were object of intensive optimization. The development, which was carried out in collaboration with the English company Electron Tubes, led to the following design [Pan04]:

- Hemispherical photocathode (PhC) and entrance window providing the following advantages:

    - Minimized time jitter (the primary electron's pathlength to the first dynode is nearly constant all over the PhC)

    - 20 %-enhanced quantum efficiency (QE) due to the possibility of double crossing

- The PhC is made of a bialkali compound with enhanced green sensitivity (wavelength range 330 nm $-$ 470 nm) to better match the Cherenkov light wavelength distribution. It was taken care, that the output aperture of the Winston cones matches the sensitive area of the PhC.

- Coating of the hemispherical entrance window with a light scattering laquer doped with a wavelength shifter. In this way the peak-QE could be increased from 25 % to 30 %. See also figure 3.9 .

- 6-stage dynode system in circular focused configuration giving the following properties (in contrast to normally 10 stages):

    - Low gain ($\lesssim 2 \cdot 10^4$), which is necessary to keep the night sky background (NSB) induced anode current small. The low gain of the PMT is compensated by a succeeding transimpedance amplifier.

    - Good interdynode e^--collection efficiencies and low interdynode time spread. Together with the hemispherically shaped PhC, signals with short rise time and small FWHM (1 $-$ 1.2 ns) are obtained, which are absolutely necessary requirements for a further suppression of NSB and for a functioning of the next neighbor (NN) trigger logic.

    - An adjustment of the interdynode voltages was done in order to combine a low afterpulsing probability with a still resonable single photoelectron (phe) response. Therefore the simple and robust F-factor method can be used in calibration (see section 3.6).

In order to reduce the camera weight and size, almost all electronics, readout and trigger parts are placed in a DAQ-building, which serves also as control building. The connection from the camera to this building is done via VCSEL drivers feeding optical fibers at the camera site and photodiodes reading out the optical fibers in the DAQ building.

Figure 3.9: The camera of the MAGIC telescope in construction phase. The photomuli-pliers are partly equipped with Winston cones. Where sun-light is illuminating the PMTs the coating becomes visible.

3.5   Camera-external signal processing

The separation of the readout and trigger electronics from the telescope camera by in-troducing an optical signal transport to an about 100 m distant DAQ building has the following advantages:

- Smaller camera of less weight, so that a better telescope mobility (less inertia) is granted. Apart from that, one obtains a better stability of the camera position (less oscillations).

- Electronics situated in a place with generous dimensions, which allows easy main-tenance, exchange or upgrade of electronic parts. Furthermore, electronic noise and the effect of cross-talk is reduced, and there are less heat dissipation problems (a cooling system is comparably easy to realize).

The obvious disadvantage of long cable ways (162 m, [B$^+$04]), which would introduce a significant distortion, spread and attenuation of the analogue signals (in the case of using standard copper coaxial cables), is met by introducing optical fibers as signal transmitters. Figure 3.10 shows the signal flow:
The PMT analogue signals are amplified by a transimpedance amplifier and fed into the transmitter board by means of (short) coaxial cables. On the transmitter boards there are drivers converting the voltage into a current signal, which is supplied to the Vertical Cavity Surface Emitting Lasers (VCSELs). These devices produce a light signal proportional to the electric current and feed this signal into optical fibers.
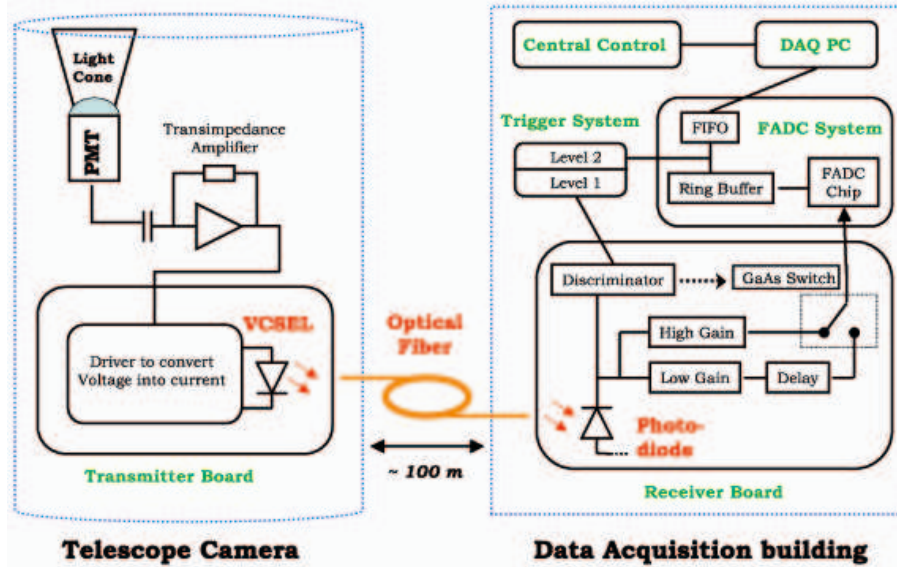
Figure 3.10: The signal flow from the camera to the DAQ building. Taken from [Pan04].

The transportation through optical fibers eliminates old problems with the transmission of normal coaxial copper cables:

- Low attenuation of 0.3 dB/100 m @ 500 MHz (in contrast to 24 dB/100 m @200 MHz for coaxial copper cables) [Pan04]

- Reduced space and weight: A cable containing 72 multimode graded index fibers has a diameter of 16 mm (a single copper cable: $\varnothing = 5\,\text{mm}$) and has about 10-times less weight than a standard coaxial copper cable [Pan04].

- No cross talk between fibres, insensitivity to electromagnetic inference.

Gain instabilities and noise of the VCSELs could be reduced below the statistical fluctuations of a PMT [Pan04]. Photodiodes on the receiver boards inside the DAQ building transform the optical into electrical signals. The succeeding signal processing is AC coupled, so that only the pre-amplified fast PMT signals are further amplified. Now the signal flow splits into two branches:
One branch leads over a discriminator (acting as 0-trigger level) to a GaAs-switch and to the level-1 and 2 trigger stages. After the signal is stretched (to a FWHM of 6 ns) in order to meet the 300 MHz-FADC sampling requirements (6 ns leads to at least 3-4 sampling points), the other branch is again split into a 50 ns-delayed low gain and an about 10-times amplified high gain branch. The previously mentioned GaAs-switch activates the low-gain recording (with 50ns delay right after the high-gain was recorded) for signals close to high gain saturation. In this way the 8-bit high gain $(0 - (2^8 - 1) = 255$ FADC counts range) is enhanced to an overall dynamic range of 11 bit (10x255 equals ca 11 bit). The equipment for signal transmission is as follows:

- 36 transmitter boards containing 18 VCSELs (16 are actually used, 36x16=576), 25 are used for the inner and 11 for the outer pixels.

- 72 receiver boards, each containing 8 channels for optical-to-digital signal conversion (72x8=576)

The MAGIC camera runs in a self triggered mode, 325 innermost pixels (out of 397 inner pixels) are participating in the trigger and set up the trigger area. This area is divided into 19 overlapping macro cells of 36 pixels each (see figure 3.11).

The trigger logic is divided into different levels:

- *Level-0*: The PMT analogue pulse, which is preamplified and transmitted via optical fibers and reconverted into an electrical signal at the receiver board, has to exceed a discriminator threshold before reaching the other trigger levels. This discriminator threshold is set by a 8 bit DAC (and can be remotely controlled by the central control PC), so that a 4mV amplitude at the input of the transmitter board (where the VCSEL units are located) corresponds to 36 DAC units. The discriminator response (in the case that the treshold is exceeded) is a rectangular shaped pulse with an (adjustable) width of 6 ns.
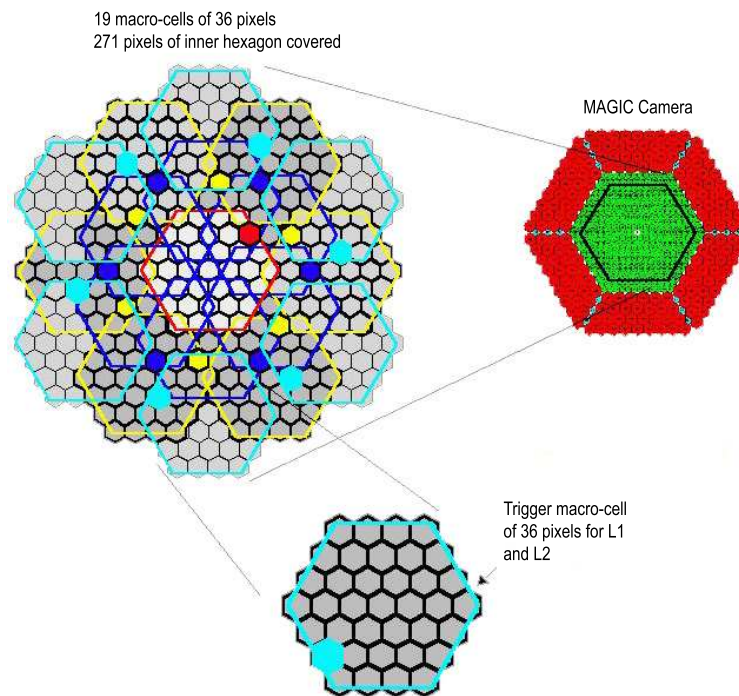
Figure 3.11: Scheme of the trigger zone of the MAGIC camera, which is divided into macro cells.

- *Level-1*: The above mentioned macrocells are realized by 19 level-1 trigger boards. In each macrocell a search for time correlated, closed packed 4NN groups (4NN = 4

next neighbour) is done. The time window is adjustable and also the NN logic (2-, 3-, 4- or 5NN) can be changed remotely by the central control PC. In the case that at least one 4NN group was found (standard setting), the last trigger stage follows.

- *Level-2*: It is foreseen that a digital analysis can be done. Therefore each macrocell (corresponding to a so called SMART module) is divided into 3x12 pixel regions (called LUT, Look Up Table). Successive stages of the digital analysis allow to impose cuts on the event topology in a tree-like structure (using the mentioned LUTs). This way a pattern recognition analysis can be done to provide an online g/h-separation and suppress hadron triggers already at the 'source'. At present the level-2 trigger is not performing a digital analysis, since studies are still going on. The second trigger level also contains prescalers in order to reduce the DAQ rate to a level < 1 kHz. Finally the DAQ is enabled by means of a level-2 trigger signal arriving at the digital modules of the FADC boards.

Thus, in normal operation the trigger is activted via the level-0 (discriminator threshold), level-1 (search for time correlated, closed packed 4NN pixel groups) and level-2 (optional prescaling, default is no prescaling) steps.

The DAQ-sytem consists of 18 crates with 4 FADC boards respectively, supplying 8 FADC channels (18x4x8=576). One channel corresponds to an 8 Bit 500 MHz bandwidth FADC chip (providing an FADC signal amplitude range from 0 to $2^8 - 1 = 255$) continuously digitizing at 300 MSamples/sec and writing the information to a 32 kByte ringbuffer. When a level-2 trigger signal arrives, the digitizing is interrupted and the readout proceeds as follows:

- 30 time slices of 1 Byte are written to a 512 kByte FIFO with a maximum rate of 80 MB/sec (this corresponds to a dead time of $< 1 \, \mu$s or $< 1 \, \%$)

- Time and trigger information are recorded by the digital modules

This information is then read out and stored in a raw data format on a RAID0 disk system at a speed of up to 20 MB/sec. In this way a maximum of 800 GB raw data per night can be processed [Cor04].

3.6 The calibration of the MAGIC telescope

In order to understand the telescope's response to a certain amount of photons hitting the camera, a calibration system is necessary. This calibration system consists of a light source, which in principle simulates typical Cherenkov light flashes, and detectors determining the number of photons emitted by this light source. The individual PMTs are then calibrated to this measured light flux. The light source is realized by a box containing 16 pulser boards equipped with LEDs of 3 different types corresponding to the emission of 3 different monochromatic light flashes (wavelengths 370 nm, 460 nm and 520 nm with a duration of 3-4 ns FWHM emitting $10^8 - 10^{10}$ photons/sr [Cor04]). The pulser box is located on the reflector at 17 m distance from the camera (see figure 3.12). A frosted glass
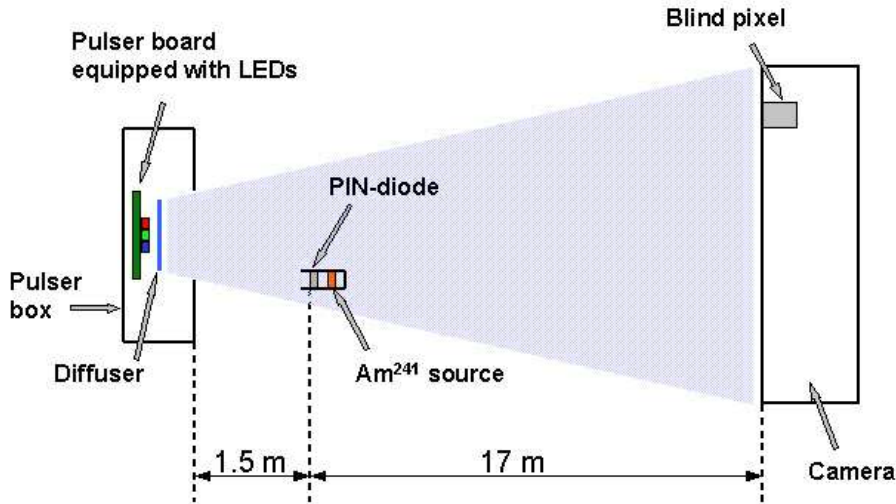
Figure 3.12: Scheme of the calibration system of the MAGIC camera.

plate in front of the pulser box acts as diffusor, so that the light flashes leaving the pulser box's window uniformly illuminate the camera (the uniformity is within 3 % [Pan04]).

There are 2 independent systems, which are foreseen to provide an absolute calibration of the detector signal chain from the Winston-Cone-PMT unit up to the FADC sytem and DAQ. They are called 'blind pixel' and 'PIN diode' method. Both methods require a measurement of the mean number of photoelectrons deposited in the detector (the blind pixel or the PIN diode respectively). The knowledge of the LED's light spectrum (usually monochromatic), the system's geometry (solid angle diminishing the light flux) and the detector's QE allows one to determine the number of photons emitted by the LED pulser. The two different methods are described as follows:

- *'Blind Pixel' method*
  A special pixel in the outermost ring of the camera is designed to measure single phe spectra. Therefore the LED flashes are attenuated by means of a $1 \, \text{cm}^2$-diaphragm and a filter transmitting 1 % of the lightflux. The single phe spectrum can be described by a sum of Gaussians with Poissonian amplitudes:

$$f(x) = \sum_{k=0}^{N} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{1}{\sqrt{2\pi}\sigma_k} e^{\frac{(x-\mu_k)^2}{2\sigma_k^2}} \tag{3.1}$$

  Here, x is the FADC-channel (the 'charge'), which behaves like a random number. μ_k and σ_k are the mean and sigma of Gaussian number k, respectively. k denotes the number of photoelectrons (phe) and λ is the mean value of the number of phe. A fit of the function 3.1 yields the mean number of phe. Taking into account the known QE and the geometry of the setup leads to a determination of the photon flux at the camera.

- *'PIN Diode' method*
  A PIN diode located at a distance of 1.5 m from the pulser box (see figure 3.12)

measures the photon flux in another way: Firstly this diode is calibrated by means of a Am^{241}-source emitting γ-rays at an energy of 59.95 keV. This produces a well known number of phe, namely 16570 ± 50 [Pan04]. Thus, the output current/voltage can be related to the number of phe, if later the pulser is used. Again by knowing the QE and the geometrical setup one can deduce the photon flux at the camera.

The calibration is then finally done by relating the photon flux to the recorded pixel charge (in FADC counts).

At the time of writing this thesis the standard calibration procedure was the simple and more robust 'excess-noise factor'-calibration [GS02], or short 'F-factor'-method. Here, in a first step, the mean number of photoelectrons is determined pixel by pixel from the relation

$$\overline{n_{phe}} = F^2 \frac{(A-P)^2}{\sigma_A^2 - \sigma_P^2} \qquad (3.2)$$

where A denotes the charge output of the PMT and P is the pedestal. The corresponding variances must be subtracted from each other as shown. The F-factor takes into account fluctuations in the PMT gain and leads to the given equation, which can be interpreted as 'F-factor corrected' Poissonian relation. The gain fluctuations σ_G and the mean gain $\overline{G}$ are related to F by [Mir00]:

$$F = \sqrt{1 + \left(\frac{\sigma_G}{\overline{G}}\right)^2} \qquad (3.3)$$

The calibration now proceeds as follows:

- *Inner pixels*
  The number n_{phe} of phe is determined pixel by pixel from the charge distribution according to equation 3.2. For each pixel the conversion from charge (FADC counts) to the average number $\overline{n_{phe}}$ of phe (over all inner pixels) is calculated.

  Thus, the F-factor method provides a calibration in phe, and by relating the individual pixel charge to $\overline{n_{phe}}$ (and not n_{phe}) one additionally achieves a 'flat-fielding'. 'flat-fielding' means, that each pixel gives the same response (in phe) to the uniform LED photon flux. The absolute photon flux can be calculated from $\overline{n_{phe}}$ and a mean QE.

- *Outer pixels*
  The outer pixels are 'calibrated' to yield a number of phe equal to 4 times the mean number of phe of the inner pixels (so they are calibrated with respect to the inner ones). This is a relative calibration in units of equivalent phe, which is actually a calibration to photons (the outer pixels are weighted according to their 4 times larger area, which is true for the incident photon flux).

There are currently two calibration procedures. One 'standard' calibration before the beginning of data taking, and a so called interleaved calibration procedure (LEDs fired

at a frequency of 50 Hz), which is run during the data taking and which allows one to monitor the stability of the electronics and, if necessary, correct the 'standard' calibration.

For a final absolute calibration local muons (which produce ring- or arc-like images) can be used, which take into account also the conditions of the close-by atmosphere. Here one relates the number of photons radiated from the muon to the measured charge. The photon flux from the muon depends on its energy (directly related to the radius of the muon ring) and the refractive index of the air as well as the geometry (inclination and impact of the muon) and can be calculated or taken from MC simulations.

Also the tuning of some parameters in the Monte Carlo can be done by means of muon rings. In this way the mirror's reflectivity and point spread function (PSF) can be both determined and inserted into the MC simulation. PSF and reflectivity especially influence the *width* and *size* distributions respectively (see next chapter), which are very important for the g/h-separation and the energy estimation.

Chapter 4

The Data Analysis Method

The analysis of raw data taken by the MAGIC Telescope proceeds in steps, which can be described as follows:

- Merging and pre-processing of raw data (the so-called 'merpping', data with information from different subsystems is merged and written in ROOT format, see [BR06] and [BWM03, Bre06])

- Calibration

- Signal Extraction and Image Restoration (restoration of the image parts, which contain significant signal information, explanations follow)

- Final Image Cleaning and Calculation of Image Parameters

The analysis steps, which follow afterwards and which involve Monte Carlo data, will be addressed later. Here, the above listed steps are discussed. Firstly, the signal extraction will be explained since it is used both in the calibration and the image restoration procedures.

The raw data provided by the MAGIC telescope consist of the counts in 30 FADC slices for each pixel. The herein contained signals must be searched for and extracted, which results in a 'charge'. This extraction is a method-dependent integration over the counts in the FADC slices. The charge has a certain variance σ^2_{charge}, which is caused by three main contributions:

$$\sigma^2_{charge} = \sigma^2_{NSB} + \sigma^2_{el} + \sigma^2_{extr} \tag{4.1}$$

σ^2_{NSB} is the noise introduced by the night sky background (NSB), which is just background light mainly due to (scattered) starlight. σ^2_{el} is the electronic noise as produced by the whole electronic chain. σ^2_{extr}, finally, denotes the extractor variance, which is proportional to the charge and which depends on the extraction method (see [SHK$^+$05b, SHK$^+$05a, B$^+$05]).

The signal extraction using the spline method will be discussed furthermore due to its simplicity and since it is among the best extractors with respect to the achievable charge resolution [SHK$^+$05b, B$^+$05]. The 'spline method' is a cubic spline interpolation of the

FADC counts, which is done prior to the signal extraction for all slices. The extraction then proceeds as integration over the relevant slices keeping σ^2_{charge} small (see below).

σ^2_{NSB} is the noise contribution, which is dominating for small and medium charges. In order to keep σ^2_{charge} as small as possible, one therefore should choose a small integration window of 2 slices [SHK+05a]. This is done if dealing with the HI-gain signal in observational data.

For relatively large charges ($\gtrsim 50-100$ phe), the intrinsic extractor error is dominating over the NSB contribution to the charge variance, and one should use a longer integration window of 6 or 7 slices [SHK+05b]. This is done if dealing with the LO-gain signal in observational data and if dealing with calibration runs.

For each pixel, the signal is extracted using the spline in the following way:

- *Signal Extraction in Calibration*

  - *Calibration run*
    The maximum of the spline is searched for within the first 15 slices (high-gain range). The signal is extracted by integrating over the spline using a window of 6 slices (2 slices to the left and 4 slices to the right of the maximum in order to account for signals with tails). Besides, it shall be pointed out, that an LED-blue or LED-UV run is preferable, since it involves comparably low fluctuations due to the high PMT QE for this spectral region.

  - *Pedestal run*
   Pedestal and pedestal RMS are extracted by integrating over the first 6 slices (fixed window position).

- *Signal Extraction in 'Image Restoration'*
  A spline-maximum search is done taking into account all 30 slices. If the maximum is found inside the high-gain range (slices 1-15), and if furthermore the high gain signal amplitude exceeds a certain threshold (typically 100 dc (digital counts)), a repeated signal search restricted to the low-gain range (slices 16-30) is performed. The signal extraction involves a spline-integration using a 2-slice window (symmetrically around the peak) for the high-gain and a 7 slice window (2 slices left from the maximum and 5 slices to the right) for the low-gain signal.

  The extraction of pedestals and pedestal RMS is done by integrating over the first two 'pre-historic' bins (slices 1 and 2), if no signal is present. Looping over 1000 events (corresponding to 4 sec taking a rate of ca. 250 Hz) yields sufficient statistics together with the ability of reacting on fast changes (rather electronic noise than e.g. changes in weather conditions or sky field rotation, which occur on much larger time scales of the order of minutes).

4.1   Calibration

The calibration procedure determines the charge-to-phe conversion factors as well as the constant (between pixel) time delays taking one pixel as reference. Furthermore, the
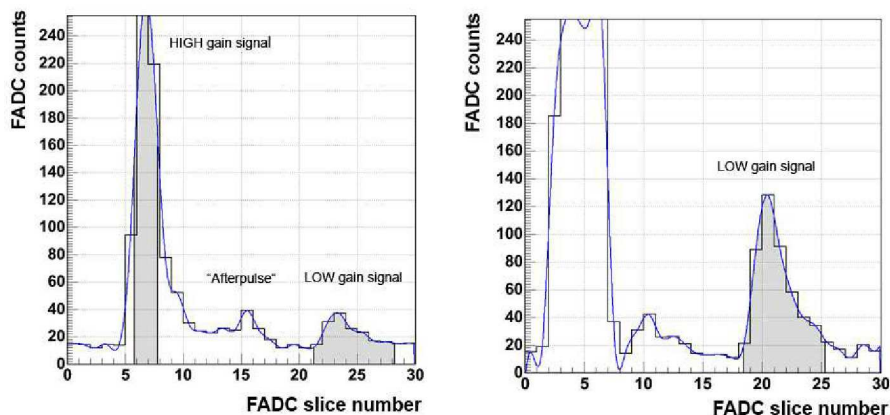
Figure 4.1: Histogram of FADC slices containing a clear signal. The histogram on the right shows a saturated high-gain, the amplitude clearly exceeds 255 dc. Taking into account the known hardware time delay of the low gain branch (about 55 ns) the 'afterpulse' cannot be mixed up with the low gain signal.

pixels are checked to work reliably (so-called dead pixel flag) and, if necessary, 'switched off'. The charge and the arrival time of these pixels will be interpolated from neighboring core pixels, when calculating image parameters. The actual calibration (calculation of conversion factors $C_{FADC \to phe}$) is done by means of the F-factor method (see section 3.6, page 55).

There is a further calibration, which is used in a kind of correction procedure, after applying the just described standard calibration:

- *Interleaved Calibration*
  The so-called interleaved calibration is done by means of dedicated runs (LED flashes), which are continuously taken during normal source observation at a rate of 50 Hz. By comparing the conversion factors from these data with those obtained in the standard calibration, one continuously obtains correction factors, which account for fast changes in the signal transmission chain (especially the VCSEL and the optical link system may introduce rapid effective gain changes). These correction factors are calculated and applied at a later analysis step, when the image parameters are determined.

- *High-to-Low Gain Calibration*
  The High-to-Low gain calibration is done from the data itself, there are no dedicated runs. Therefore, signals appearing in the (non-saturated) high- and, simultaneously, low-gain channels are taken and the high-to-low gain conversion factors as well as the corresponding time delays are calculated.

4.2   Signal Extraction and Image Restoration

After the standard calibration constants have been determined, the further off-line analysis proceeds with the 'image restoration'. A detailed description of this procedure, which can also be regarded as image cleaning method, together with the involved extractor properties can be found in [SHK$^+$05b].
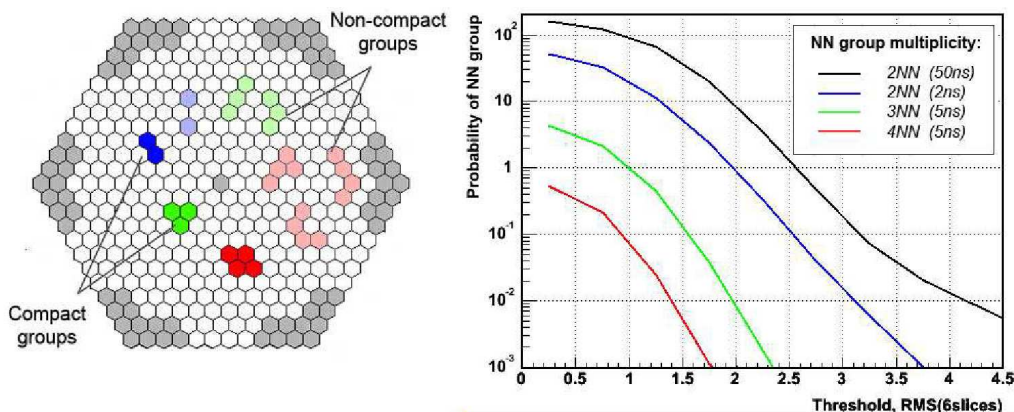
Figure 4.2: Left: Examples for 4-, 3- and 2NN groups inside the inner part of the MAGIC camera. Pixels outside the hardware-trigger region are drawn grey-shaded. **Right:** Multiplicity of 4-, 3- and 2NN groups appearing in pedestal data (time window given in brackets) as a function of the image cleaning threshold. The threshold is expressed in units of pedestal RMS σ_P for a 6-slice-integration. Since there is a variety of different extractor methods and integration ranges [B$^+$05], the 6-slice-integration window is taken as reference. In the case of Crab Nebula observations, the NSB level corresponds to $\sigma_P = 2.72$ phe [SHK$^+$05b].

The basic idea of 'image restoration' is to 'restore' the original 4NN trigger (see section 3) and successively build up the image out of close packed next neighbour (NN) groups. In a 'close packed' or 'compact' NN group of 2-4 pixels, each pixel has one (2NN) or two (3NN and 4NN) nearest neighbors (see the plot on the left in figure 4.2).

The standard procedure, which is implemented in the Mars (MAGIC Analysis and Reconstruction Software) package [BWM03], uses a 2NN pixel search without restricted FADC range (the two pixels, which make up the 2NN group, can have a corresponding arrival time difference up to 15 FADC slices or, equivalently, 50 ns). In order to suppress NSB noise signals, it is advisable to impose a tight time coincidence onto the NN groups, which leads to a *time-correlated close-packed NN search*.

The initial trigger is a 4-pixel close-packed next neighbour group (each pixel exceeding a certain discriminator threshold, see section 3) within a time window of 6 ns. Since the information about the initial trigger is not recorded and due to a limited time resolution ($\sigma_{reflector} = 400$ ps and $\sigma_{DAQ} \approx 600$ ps thus $\sigma_{total} = \sqrt{\sigma_{reflector}^2 + \sigma_{DAQ}^2} \approx 700$ ps not taking into account problems for low signal amplitudes , see [SHK$^+$05b]) the original trigger can

NN Multiplicity	Signal Threshold	Coincidence window
4	1.6 σ (4.4 phe)	$\pm$ 2.5 ns
3	2.0 σ (5.4 phe)	$\pm$ 2.5 ns
2	3.0 σ (8.2 phe)	2 ns
2 ('ring pixel')	1.1 σ (3.0 phe)	3.0 ns

Table 4.1: Signal thresholds and time-coincidence windows for the image restoration procedure. The thresholds are given for a Crab-like night sky background.

be lost, especially when approaching the telescope's trigger threshold. Therefore, also 2NN and 3NN groups are involved in the first step towards image restoration (see figure 4.2).

As already mentioned, the hardware trigger is a 4NN coincidence within a time window of 6 ns. For the 'image restoration' process, which acts as a kind of 'software trigger', the coincidence windows are determined from the arrival time variance σ_t^2 of a low 4 phe signal in presence of a Crab-like NSB level. Here $\sigma_t \lesssim 2\,\text{ns}$ can derived [SHK$^+$05b]. By applying the NN search procedure to pedestal data, one can determine the probabilities of catching NSB noise (see figure 4.2, right part).

In the ideal case, one should set a threshold right above the real trigger threshold. Preliminary studies have shown, that for the real hardware trigger (4NN multiplicity, time window of 6 ns) the probability of catching NSB noise is about 1 % [SHK$^+$05a]. In order to avoid an unnecessary complexity of the image cleaning method, the thresholds for 3NN and 2NN pixel groups are chosen in such a way, that roughly equal probability levels are obtained. In this way, one can derive the settings as given in table 4.1.

The whole 'image restoration' process can be expressed as follows (see also figure 4.3):

- Search for 2-, 3- and 4NN time-correlated pixel groups within the inner camera (trigger region, see figure 3.11). If no NN group is found the event is rejected.

- Search for 2-, 3- and 4NN time-correlated pixel groups in the whole camera. Now, the image core is completely 'restored'.

- Define rings of neighbours around the just restored image core (usually 2 rings are sufficient to define a smooth image boundary). The signal in these 'ring pixels' are re-extracted using a narrow peak-search window of only 1 slice. For a certain 'ring pixel' the position of this window (the expected signal position) is determined from the mean arrival time of the next-neighbour core pixels (1st ring) or the next-neighbour pixels, which have survived the succeeding cleaning (2nd ring).

This sophisticated image restoration procedure is especially suited to reduce pixelization effects, which occur close to the trigger threshold, where images consist of just a few pixels [SHK$^+$05b]. In the energy range below 100 GeV about 21 % more events are recovered in the presence of Crab Nebula like NSB [SHK$^+$05a] in comparison with a standard 2NN image cleaning. Thus, the image restoration procedure is supporting the
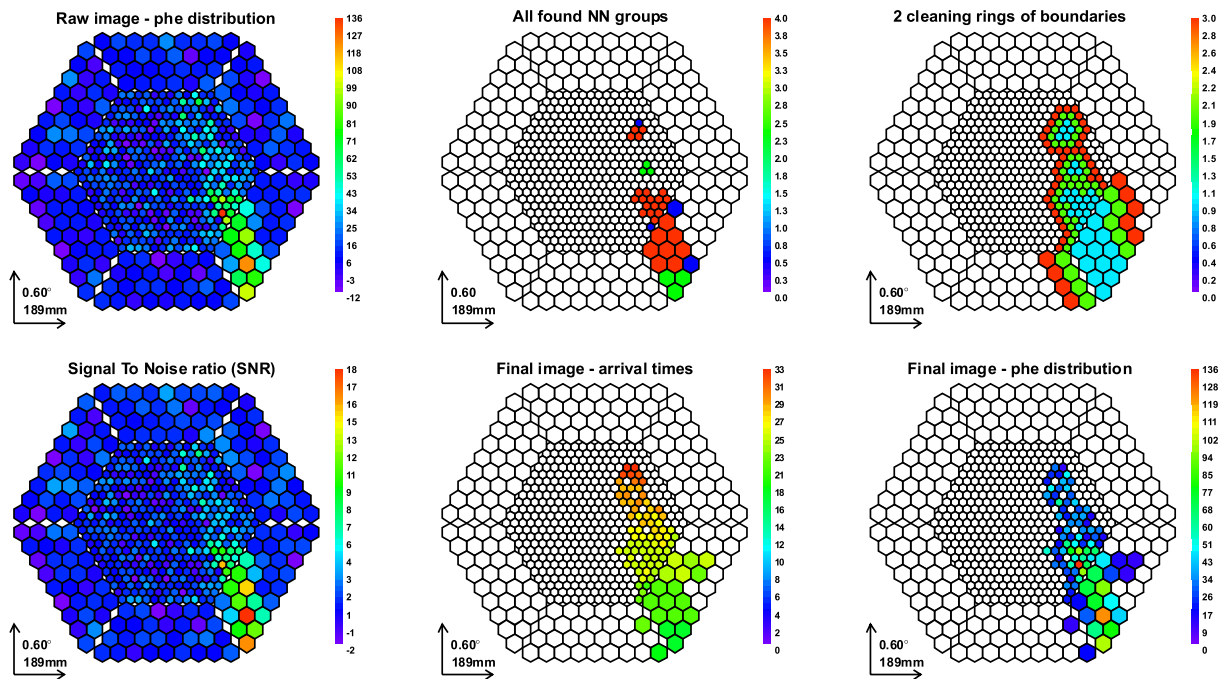
Figure 4.3: Camera displays of a cosmic event (MC simulation): **Left side:** The raw image and corresponding signal to noise ratio, the scale is in phe. **Middle top:** Found NN groups: red pixels participate in 4NN groups, green pixels are additionally found in a 3NN group search and blue ones are additionally found in a 2NN group search. **Right top:** Green and red pixels show the two rings of boundary pixels, which are arranged around the restored image core. **Middle and Right bottom:** The final image after 4-, 3- and 2NN pixel group search as well as ring cleaning. The arrival times are given in ns.

subsequent analysis, in particular advanced g/h-separation methods trying to lower the energy threshold.

A critical point, related to the agreement between observational and Monte Carlo data, is, that the simulated noise in Monte Carlo data must actually match that one in the observational data. Since NSB and electronic noise cannot be adjusted very precisely due to their complexity (e.g. one must take into account the source dependent starfield, which rotates in the camera due to the azimuth-alt telescope mounting), an 'absolute' image cleaning is additionally performed:

For a certain sky region a (mean) RMS in units of phe can be calculated. On each pixel, which survives the image restoration procedure, an absolute cleaning is additionally imposed by requiring a typically 10 % higher threshold than this mean RMS (see table 4.1 for the relation of RMS and absolute number of phe in the case of the Crab Nebula sky region). The same absolute clean levels are then used for the MC data.

4.3 Calculation of Image Parameters

The above described image restoration process provides a cleaned, high confident Cherenkov light image, in which each pixel is assigned a charge q, the NSB induced RMS and an arrival time t. The next step is the calculation of image parameters, which characterize the image and which are used in the further analysis, in particular the g/h-separation and the energy estimation.

4.3.1 Standard (Hillas) and additional parameters

The calculation of the main image parameters is performed using the standard Mars package [BWM03]. The following overview shows some of these image parameters as well as additional ones, which proved to be useful in later analysis steps (see section 2.2 for a mathematical evaluation of the standard parameters):

Source independent parameters

- Hillas Parameters

  - *width*, *length* (the two half axes of the Hillas ellipse)
  - *size* (total number of photo electrons (phe) summing up all pixels, which survive the cleaning → size of image)

- Additional parameters

  - *dens* (density defined as $\log_{10}(size)/(width \cdot length)$)
  - *leakage* (*size* in outermost pixel ring divided by *size* of image)
  - *relative island size* (*size* of image exclusive main island divided by total *size*)

The charge [phe] of pixel i is denoted by q_i. Thus, one finds:

$$size = \sum q_i \tag{4.2}$$

In the following, $\boldsymbol{r}_i$ denotes the vector, which is pointing from the source position to the center of pixel i.

Source dependent parameters

- Hillas Parameters (see also section 2.2)

  - *dist* (distance from center of image (= center of gravity, shortly *cog*) to source position, which is usually the center of the camera (but not in the so-called wobble mode, where ON and OFF data are taken simultaneously, or if there is significant mispointing)

  - *alpha* (angle between the major axis of the image and the line from the *cog* to the center of camera)

- Additional parameters

  - *Charge Asymmetry*
    Given the definitions

$$\boldsymbol{r}_q = \frac{\sum \boldsymbol{r}_i \cdot q_i}{\sum q_i} \tag{4.3}$$

$$\boldsymbol{r}_{q^2} = \frac{\sum \boldsymbol{r}_i \cdot q_i^2}{\sum q_i^2} \tag{4.4}$$

$$\boldsymbol{a}_q = \boldsymbol{r}_{q^2} - \boldsymbol{r}_q \tag{4.5}$$

the cosine of the charge asymmetry related angle and the length of the charge asymmetry vector are calculated as follows (see also figure 4.4):

$$\cos \beta_q = \cos(\angle(\boldsymbol{a}_q, \boldsymbol{r}_q)) = \frac{\boldsymbol{a}_q \cdot \boldsymbol{r}_q}{|\boldsymbol{a}_q| \cdot |\boldsymbol{r}_q|} \tag{4.6}$$

$$length_q = |\boldsymbol{a}_q| \tag{4.7}$$

The idea of calculating a charge asymmetry arises due to the asymmetric intensity profile of the image along its major axis (see plot on the left in figure 4.4).

In the g/h-separation and the energy estimation (see next chapter) the following parameters are discussed:

  * head-tail information calculated as $sign_q = \text{sign}(\cos \beta_q)$
  * length of charge asymmetry vector $length_q = |\boldsymbol{a}_q|$

– *Time Asymmetry*
In analogy to the charge asymmetry, the definitions

$$\boldsymbol{r}_t = \frac{\sum \boldsymbol{r}_i \cdot q_i \cdot t_i}{\sum q_i \cdot t_i} \tag{4.8}$$

$$\boldsymbol{r}_{t^2} = \frac{\sum \boldsymbol{r}_i \cdot q_i \cdot t_i^2}{\sum q_i \cdot t_i^2} \tag{4.9}$$

$$\boldsymbol{a}_t = \boldsymbol{r}_{t^2} - \boldsymbol{r}_t \tag{4.10}$$

lead to the following image parameters:

$$\cos \beta_t = \cos(\measuredangle(\boldsymbol{a}_t, \boldsymbol{r}_t)) = \frac{\boldsymbol{a}_t \cdot \boldsymbol{r}_t}{|\boldsymbol{a}_t| \cdot |\boldsymbol{r}_t|} \tag{4.11}$$

$$length_t = |\boldsymbol{a}_t| \tag{4.12}$$

In the energy estimation (see next chapter) the combined parameter $\text{sign}(\cos \beta_t) \cdot |\boldsymbol{a}_t|$ is discussed in some detail.
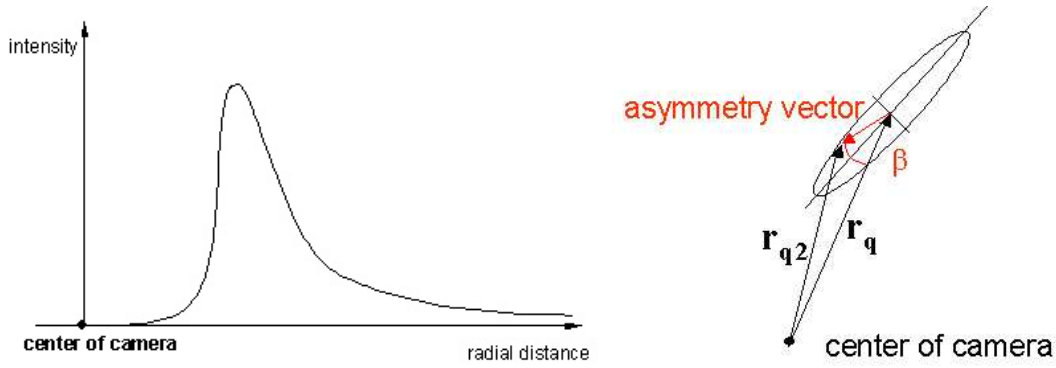
Figure 4.4: Origin of the charge asymmetry and definition of the charge asymmetry vector. **Left:** Intensity profile of the Cherenkov light image along the major image axis. The asymmetric shape (long image tail pointing away from the camera center for parallel shower and telescope axes) arises due to geometrical reasons, compare also figure 2.6. **Right:** Topview of the Cherenkov light image as described by the Hillas ellipse. $\boldsymbol{a}_q$ is drawn as red arrow and $\measuredangle(\boldsymbol{a}_q, \boldsymbol{r}_q)$ is denoted as β.

The time asymmetry arises since Cherenkov photons emitted at different stages in the shower development arrive at different times at the reflector. Here, one must consider that the Cherenkov photons emitted high in the atmosphere are 'overtaken' by the Cherenkov light emitting charged shower particles.

Therefore, at relatively small distances of the telescope to the shower axis (impact parameter $R < 120$ m, which is the position of the hump in the Cherenkov light pool), Cherenkov photons emitted at later stages of the shower development in the atmosphere

are actually arriving earlier than the photons emitted at the beginning of the shower development.

However, the total path length, determined by the charged particles' and the Cherenkov photons paths, becomes larger when the Cherenkov photons under consideration are emitted at later stages in the shower development (i.e. by electrons out of a dying particle cascade). Thus, the temporal image structure swaps over at a large distance from the shower axis due to the rather long travel path in a dense atmosphere for the photons emitted by electrons out of a dying particle cascade. This is a well-known effect, sometimes called the 'seagull effect' [Kon04].

In short, the interplay of path length and travel velocity of the electrons and the produced Cherenkov photons determines the direction of the temporal development of the camera image, which swaps at an impact parameter $R \approx 120$ m.

The arrival times t_i are calculated with respect to the trigger time, which is in MC simulations constant. However, this is not the case for real observational data. Therefore, the time asymmetry parameters will be exploited merely in an MC study of the energy estimation.

4.3.2   Image parameters related to muon rings

Besides the image parameters, a muon ring fit provides further valuable information particularly as input for an adjustment of the main MC simulation parameters *reflectivity* and *point spread function (PSF)*. The Monte Carlo simulations (muons, protons and gammas) used in this work have been produced with the standard MAGIC Monte Carlo chain. The shower simulation is based on the well-known CORSIKA code [HK02]. Detailed information in particular concerning the subsequent MAGIC Monte Carlo branches (reflector and camera simulations) can be found in [Bla04], chapter 4.

The basic parameters entering in the muon simulation are:

- *Starting point of muon track* $T_0 = 600$ g/cm^2

- *Muon energy range* 5 GeV $< E_\mu < 1$ TeV
  The spectral index is chosen in order to approximate muon flux measurements: $\alpha = -2.7$ for $E > 10$ GeV and $\alpha = -1.4$ for $E < 10$ GeV reflecting the turnover in the muon flux induced by muon energy loss and muon decay, which become dominant for muon energies below 10 GeV (see [Gai90], figure 6.1 on page 71).

- *Impact parameter* $R_\mu < 15$ m and *viewcone angle* $\delta < 1.2°$
  The muon track directions are generated within a circular cone around the fixed primary axis being parallel to the telescope axis. The half opening angle of this cone is adjusted to $1.2°$ (radius of camera). Therefore, the muon arcs in the camera are shifted by an angle $-1.2° < \Delta < 1.2°$ away from the centered position. In other words, the muon arc centers are distributed over the whole camera.

  The impact parameter R_μ and the telescope radius R_{tel} determine the muon arc-angle ϕ_μ via $\sin \frac{\phi_\mu}{2} = \frac{R_{tel}}{R_\mu}$. Thus the minimum arc-angle is $\phi_\mu = 70°$. For the calibration purpose only muon arcs with $\phi_\mu > 160°$ are taken (see [Mey05]).

The muons finally taken for a MC adjustment (i.e. MC muons as well as muonic background in observational data) are selected by applying cuts in the ring fit χ^2 and *arc length* (the ϕ-angle described above). A lookup-table basically containing muon MC simulations for different settings of reflectivity and PSF is used to adjust the MC data to real observational data by means of certain image parameters, which can be derived from a muon ring fit. These are the *arc width* (w_μ) (determined from a Gaussian fit to the radial intensity distribution, see below), the *arc length* (ϕ_μ) and the *arc radius* (ρ_μ).

Reflectivity adjustment

The *arc radius* ρ is correlated with the muon energy via the Cherenkov angle $\cos \rho = \frac{1}{n\beta}$. Therefore, the energy-*size* relation can be investigated by means of the $size(\rho)$ relation. In this way, the reflectivity (which is the main factor influencing the energy-*size* dependence, if one assumes a correctly simulated QE of the PMTs) can be checked and adjusted.

PSF adjustment

Muon rings offer the opportunity of adjusting the PSF between MC and real data by means of the *arc width* distribution w_μ. The *arc width* of a single muon image is calculated by integrating the radial intensity profiles (the integration is over the azimuthal angle ϕ with the maximum ϕ being the arc-angle) and fitting the distribution by a Gaussian. The fit range is defined by those intensities, which exceed half of the maximum intensity - a constraint, which is needed when dealing with muons selected from observational data, since these may contain hadronic signals affecting the tails of intensity profile. The lower value of the *arc width* is limited by the PSF.

Histograms of the *arc width* and *size* versus *arc radius* diagrams as discussed above are shown in chapter 6 and in the appendix: See the figures 6.2 and 6.3 for the results of the MC adjustment in the case of the analysis of Crab Nebula observational data as well as the figures A.10, A.8 and A.9 for the MC adjustment in the analysis of PG1553+113 observational data. The reader is also also referred to [MBH+03] and [Mey05] for details on the methods of reflectivity and PSF measurement by means of muon rings.

The usage of muon ring parameters helps to identify and reject pure muon and muon dominated images, and therefore, besides the mentioned MC adjustment, a possible utilization of muon ring fit related parameters for the g/h-separation arises. A simple parameter, which 'feels' the curvature of the image, proved to be useful [Sha05]. The 'muon ring center distance' d_μ is given as:

$$d_\mu = -\sin\delta \cdot (x_\mu - x_{cog}) + \cos\delta \cdot (y_\mu - y_{cog}) \qquad (4.13)$$

Here, δ is the angle between the major axis of the Hillas ellipse and the x-axis of the camera coordinate system (see section 2.2 and figure 2.7). The indices μ and *cog*, when appearing as coordinate index, refer to the muon ring center and the center of gravity, respectively. Thus, the above equation is a transformation into the coordinate system of the Hillas ellipse. The parameter d_μ is calculated as y-coordinate of the distance vector $\boldsymbol{d} = \boldsymbol{r}_\mu - \boldsymbol{r}_{cog}$ in the Hillas ellipse coordinate system (see figure 4.5).
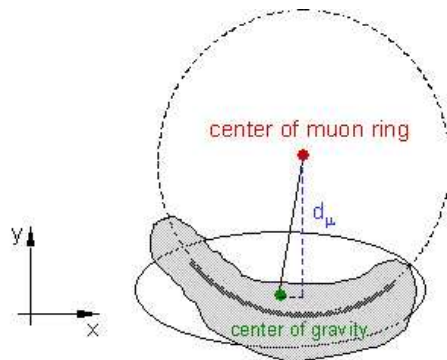
Figure 4.5: Scheme showing the d_μ-parameter. The coordinate system is chosen such, that the x-axis is parallel to the major axis of the ellipse and the y-axis is parallel to the minor axis of the ellipse. Only the y-coordinate of the distance $\boldsymbol{d} = \boldsymbol{r}_\mu - \boldsymbol{r}_{cog}$ is taken since the x-component is more dominated by fluctuations due to the (by definition) larger variance along the major ellipse axis.

Plots of all image parameters, which were discussed in this section, can be found in chapter 6.

4.4   Flux calculation and spectrum

A particle flux F is defined as the number of particles per area dA, time dt and solid angle $d\Omega$.

$$F = \frac{N}{dA \cdot dt \cdot d\Omega} \tag{4.14}$$

In the case of a point source the solid angle dependence cancels out. Thus, one obtains

$$F \;=\; \frac{N}{dA \cdot dt} \tag{4.15}$$

$$\Rightarrow \frac{dF}{dE} \;=\; \frac{dN}{dE \cdot dA \cdot dt} \tag{4.16}$$

The last equation describes the so-called differential flux $\frac{dF}{dE}$ per energy interval dE (E denotes the primary gamma energy). What is measured by a Cherenkov telescope is not directly the flux, but the number of excess events. The number of excess events per energy interval can be expressed as:

$$\frac{dN_{excess}}{dE} = \int \left(\frac{dF}{dE}\right) \cdot \varepsilon(E, \theta, R) \cdot dt \cdot 2\pi R dR \tag{4.17}$$

Here, θ denotes the zenith angle. The integration extends over the impact-ring radius R (with $2\pi R\, dR = dA$) and takes into account the total γ-efficiency ε:

$$\varepsilon = \frac{N_{selected}(E, \theta, R)}{N_{total}(E, \theta, R)} \tag{4.18}$$

ε consists of the trigger and cut efficiencies and is therefore given as

$$\varepsilon(E, \theta, R) = \varepsilon_{trigger}(E, \theta, R) \cdot \varepsilon_{cuts}(E, \theta, R) \tag{4.19}$$

The integral $\int dt$ can be expressed as sum of observation times $T_{obs}(\theta_i)$ at discrete zenith angle intervals θ_i. The measured differential excess rate $\frac{dN_{excess}}{dE}$ is then:

$$
\begin{aligned}
\frac{dN_{excess}}{dE} &= \frac{dF}{dE} \cdot \sum_{i=0}^{N} T_{obs}(\theta_i) \int \varepsilon(E, \theta_i, R) \cdot 2\pi R dR \tag{4.20} \\
&= \frac{dF}{dE} \cdot \sum_{i=0}^{N} T_{obs}(\theta_i) \cdot A_{coll}(E, \theta_i) \tag{4.21}
\end{aligned}
$$

In the last step the effective collection area $A_{coll}(E, \theta_i) = 2\pi \int \varepsilon(E, \theta_i, R) \cdot R dR$ was introduced. Together with ε the effective collection area must be calculated from MC simulations, where energy, impact parameter and zenith angle are known:

$$
\begin{aligned}
A_{coll}(E, \theta_i) &= 2\pi \int \varepsilon(E, \theta_i, R) \cdot R dR \tag{4.22} \\
&\approx \pi \sum_{j=0}^{j_{max}} \varepsilon(E, \theta_i, R_j) \cdot (R_{up,j}^2 - R_{lo,j}^2) \tag{4.23} \\
&= \pi \sum_{j=0}^{j_{max}} \frac{N_{selected}(E, \theta_i, R_j)}{N_{total}(E, \theta_i, R_j)} \cdot (R_{up,j}^2 - R_{lo,j}^2) \tag{4.24}
\end{aligned}
$$

It shall be noted here, that, instead of integrating over ring areas, the maximum impact parameter R_{max} of the gamma MC simulation multiplied with $\frac{N_{sel}}{N_{tot}}$ can be used:

$$A_{coll}(E, \theta_i) = \pi \frac{N_{selected}(E, \theta_i)}{N_{total}(E, \theta_i)} \cdot R_{max}^2 \tag{4.25}$$

Finally, the differential flux is calculated as:

$$\frac{dF}{dE} = \frac{\frac{dN_{excess}}{dE}}{\sum_i T_{obs}(\theta_i) \cdot A_{coll}(E, \theta_i)} \tag{4.26}$$

In order to calculate the error of the differential flux, one must propagate the errors $\sigma_{A_{coll}}$ of the collection area and $\sigma_{N_{ex}}$ of the number of excess events. Introducing the abbreviations

$$
\begin{aligned}
N_{selected}(E, \theta_i) &= N_{sel,i} \tag{4.27} \\
N_{rejected}(E, \theta_i) &= N_{rej,i} \tag{4.28} \\
N_{total}(E, \theta_i) &= N_i \tag{4.29} \\
N_i &= N_{sel,i} + N_{rej,i} \tag{4.30}
\end{aligned}
$$

one finds following binomial probabilities

$$p_i \;\; = \;\; \frac{N_{sel,i}}{N_i} \tag{4.31}$$

$$q_i \;\; = \;\; \frac{N_{rej,i}}{N_i} \tag{4.32}$$

and the binomial variance

$$\sigma^2_{sel,i} = N_i \, p_i \, q_i \tag{4.33}$$

Finally the error of the collection area (equation 4.25) can be calculated as

$$\sigma_{A_{coll}} = \pi \sqrt{\frac{p_i \cdot q_i}{N_i}} \cdot R^2_{max} \tag{4.34}$$

However, for a strict treatment of the problem, the correct distribution is a Poisson law for the total number N_i of events, and a binomial law for $N_{sel,i}$ and $N_{rej,i}$ conditional on N_i. Nevertheless, the result is the same (see section 4.1.3, example iii in [EDJ$^+$71]). The calculation of $\sigma_{N_{ex}}$ is addressed in the next section.

4.5   Determination of the number of excess events and its significance

The number N_{ex} of excess events is determined by means of a so-called *alpha* histogram (see figure 4.6 as well as sections 4.3 and 2.2 for the definition of *alpha*). The parameter *alpha* describes the image orientation in the camera, and is the strongest g/h-discriminator: 'Gamma images' point towards the camera source position whereas 'hadronic images' show a random orientation. Therefore, a gamma-signal appears as Gaussian-like peak around *alpha* = 0 on top of a rather constant hadronic background, when filling a histogram of *alpha* taking only events, which survive the g/h-separation (see the diagram on the left in figure 4.6). The g/h-separation method is principally based on cuts in image parameters in order to reject a maximum fraction of hadronic events while keeping gamma events (see next chapter for more details on g/h-separation). By means of an *alpha* histogram the existence of a signal can be proven and its strength can be determined. If *alpha* would already be taken as a parameter for the g/h-separation the hadronic background could not be evaluated with the standard '*alpha* histogram method': Cuts in *alpha* 'destroy' the rather constant background at $|alpha| \gtrsim 20°$ and make a ON-OFF normalization difficult. The image parameters used for the g/h-separation method are not or just weakly correlated with *alpha*.

The distribution of *alpha* may be represented by a function $f(alpha)$. Then, the distribution of $|alpha|$ can be expressed by a function $g(|alpha|)$:

$$g(|alpha|) \;\; = \;\; f(|alpha|) + f(-|alpha|) \tag{4.35}$$

$$g^{(n)}(|alpha|) \;\; = \;\; f^{(n)}(|alpha|) + (-1)^n \cdot f^{(n)}(-|alpha|) \tag{4.36}$$

Requesting continuity at $alpha = 0$ for all derivatives $f^{(n)}$, i.e.

$$\lim_{alpha \to 0^+} f^{(n)}(alpha) = \lim_{alpha \to 0^-} f^{(n)}(alpha)$$

leads to vanishing derivatives $g^{(n)}(|alpha|)$ in the case of odd n. Thus, when $g(|alpha|)$ is represented by a polynomial, only terms of even degree are considered. For the analysis of data taken during Crab Nebula OFF as well as PG1553+113 OFF observations (see chapters 6 and 7), the OFF data $alpha$ histograms could be well fitted by a fourth-degree polynomial in the range $0^{\circ} < |alpha| < 70^{\circ}$.

If not stated otherwise, the $alpha$ histogram or diagram always shows the distribution $g(|alpha|)$ introduced above and $|alpha|$ will be denoted as α in the following equations.

An integration or summing over the first bins (signal region corresponding to $\alpha < 15^{\circ} - 20^{\circ}$) of the $alpha$ histograms of ON and OFF data yields the numbers N_{ON} and N_{OFF} respectively. The ON data are taken during 'on source' observations and contain the gamma signature (the signal peak). The OFF data are taken 'off source' observing a close-by sky region. The number N_{ex} of excess events and its significance S are defined as:

$$N_{ex} = N_{ON} - N_{OFF} \tag{4.37}$$

$$S = \frac{N_{ex}}{\sigma_{N_{ex}}} \tag{4.38}$$

In the simplest case of simultaneous ON and OFF observations one obtains

$$S = \frac{N_{ON} - N_{OFF}}{\sqrt{N_{ON} + N_{OFF}}} \tag{4.39}$$

$$= \frac{N_{ex}}{\sqrt{N_{ex} + 2\,N_{OFF}}} \tag{4.40}$$

However, for the data analyzed in this work ON and OFF data were taken non-simultaneously. The OFF data $alpha$ histogram $h_{OFF}(\alpha)$ is then scaled to the background level of the ON data $alpha$ histogram $h_{ON}(\alpha)$ by a factor

$$c = \sum_{\alpha > 20^{\circ}} h_{ON}(\alpha) / \sum_{\alpha > 20^{\circ}} h_{OFF}(\alpha)$$

Here, the sums extend over the background region in both OFF and ON $alpha$ histograms, which usually can safely be taken as $\alpha > 20^{\circ}$. Since the ratio between ON and OFF data cannot be precisely determined from the corresponding observation times (there are uncertainties due to slightly changing data taking rates), this scaling is unavoidable.

One needs high OFF data statistics to judge the background level of the $alpha$ histogram, particularly in the case of weak sources. The 'off source' observations are not contemporaneously, there is a possible shift of days. It is therefore important to check the consistency of ON and OFF data $alpha$ histograms: All $alpha$ histograms presented

in this work show the ON and the (scaled) OFF data *alpha* histograms super-imposed as well as the polynomial fit to the scaled OFF data *alpha* histogram. The polynomial describing the scaled OFF data *alpha* histogram and its integral may be denoted by $p(\alpha)$ and $P(\alpha)$ respectively.
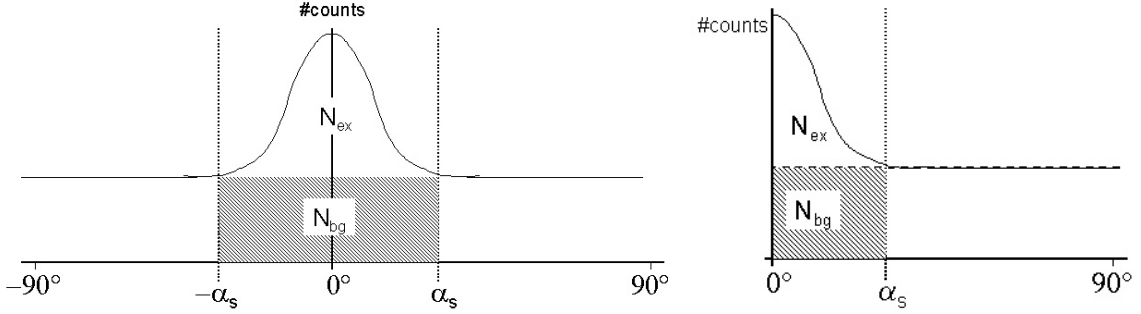
Figure 4.6: These schematic plots show the calculation of N_{ex} and N_{bg}. The horizontal, dashed line, which represents the background level, can be assumed to be constant only in the simplest case. A more appropriate description is a polynomial of degree 2 or 4. **Left:** *alpha* histogram for the full range of $-90^\circ < alpha < 90^\circ$ **Right:** *alpha* histogram showing $|alpha|$. The graphic is not to scale: The numbers N_{bg} and N_{ex} in the left and right histograms are identical.

The number N_{bg} can now be determined by integrating over the polynomial $p(\alpha)$ from 0 to α_s, where α_s corresponds to the limit of the region of significant signal.

$$p(\alpha) = a + b \cdot \alpha^2 + c \cdot \alpha^4 \tag{4.41}$$

$$P(\alpha) = a \cdot \alpha + b \cdot \frac{\alpha^3}{3} + c \cdot \frac{\alpha^5}{5} \tag{4.42}$$

Let σ_a^2, σ_b^2 and σ_c^2 denote the variances of the polynomial coefficients a, b and c. Furthermore, σ_{ab}, σ_{ac} and σ_{bc} are the covariances of a and b, a and c as well as b and c respectively. Introducing the abbreviations

$$\alpha_{s3} = \frac{\alpha_s^3}{3} \tag{4.43}$$

$$\alpha_{s5} = \frac{\alpha_s^5}{5} \tag{4.44}$$

the number N_{bg} of background events calculated as polynomial integral, and its variance $\sigma_{N_{bg}}^2$ reads:

$$N_{bg} = a \cdot \alpha_s + b \cdot \alpha_{s3} + c \cdot \alpha_{s5} \tag{4.45}$$

$$\sigma^2_{N_{bg}} = \begin{pmatrix} \alpha_s & \alpha_{s3} & \alpha_{s5} \end{pmatrix} \cdot \begin{pmatrix} \sigma^2_a & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma^2_b & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma^2_c \end{pmatrix} \cdot \begin{pmatrix} \alpha_s \\ \alpha_{s3} \\ \alpha_{s5} \end{pmatrix} \qquad (4.46)$$

Thus, the variance $\sigma^2_{N_{bg}}$ can be calculated as

$$\sigma^2_{N_{bg}} = \sigma^2_a \cdot \alpha^2_s + \sigma^2_b \cdot \alpha^2_{s3} + \sigma^2_c \cdot \alpha^2_{s5} \; + \; 2 \cdot \sigma_{ab} \cdot \alpha_{s3} \cdot \alpha_s \; + \; 2\sigma_{ac} \cdot \alpha_{s5} \cdot \alpha_s \; + \; 2 \cdot \sigma_{bc} \cdot \alpha_{s5} \cdot \alpha_{s3} \quad (4.47)$$

There are different ways of calculating the significance of a signal based on the numbers N_{ex} and N_{bg} as determined by means of an *alpha* histogram. In the following the approach of [LM83] will be described and generalized in order to make the usage of N_{bg} and $\sigma^2_{N_{bg}}$ as determined by a polynomial fit possible.

The number N_{bg} of background events can be determined directly as $N_{bg} = c \cdot N_{OFF}$. The variance $\sigma^2_{N_{bg}}$ is then given as $\sigma^2_{N_{bg}} = c^2 \cdot N_{OFF}$ taking into account a Poissonian error of N_{OFF}. This leads to the significance calculated according to [LM83], equation 5:

$$S = \frac{N_{ON} - c \cdot N_{OFF}}{\sqrt{N_{ON} + c^2 \cdot N_{OFF}}} \qquad (4.48)$$

Calculating the significance from a likelihood ratio (for details see [LM83], equation 17) leads to a normally distributed significance:

$$S = \sqrt{2} \left(N_{ON} \ln \left[\frac{1+c}{c} \frac{N_{ON}}{N_{ON} + N_{OFF}} \right] + N_{OFF} \ln \left[(1+c) \frac{N_{OFF}}{N_{ON} + N_{OFF}} \right] \right)^{\frac{1}{2}} \qquad (4.49)$$

If N_{bg} and $\sigma^2_{N_{bg}}$ are determined from a polynomial fit according to equations 4.45 and 4.47, modifications are needed. Equation 4.48 can easily be generalized in the following way.

$$S = \frac{N_{ON} - N_{bg}}{\sqrt{\sigma^2_{N_{ON}} + \sigma^2_{N_{bg}}}} \qquad (4.50)$$

However, in order to evaluate equation 4.49 in the case of a polynomial fit, one must introduce effective values of c and N_{OFF} derived from the relations $N_{bg} = c \cdot N_{OFF}$ and $\sigma^2_{Nbg} = c^2 \cdot N_{OFF}$ [Wit06].

$$c = \frac{\sigma^2_{N_{bg}}}{N_{bg}} \qquad (4.51)$$

$$N_{OFF} = \frac{N^2_{bg}}{\sigma^2_{N_{bg}}} \qquad (4.52)$$

Apart from the number of excess events together with its significance as determined from an *alpha* histogram a further 'standard histogram' is useful in the analysis of a gamma-ray source:

The so-called *false source plot method* is used to derive a map of the observed sky region in terms of excess events or significance: At first a grid of source positions covering the whole camera area is defined. Then for each of the possible source positions in the grid the ON and OFF data events are reprocessed in order to recalculate *alpha* and subsequently N_{ex}, N_{bg} and S. Thus, to each point in the mentioned grid values N_{ex} and S can be assigned. In this way one obtains a significance sky-map of the source region within the field of view of the telescope.

The false source plot is helpful for the analysis of extended sources and/or for the investigation of the telescope's pointing accuracy.

Chapter 5

The Random Forest Method

Gammas make up only a tiny fraction of cosmic rays (as small as $0.1\,\%$, see section 1.2). Therefore, even after the trigger, IACTs provide raw data with a signal-to-noise ratio below $1\,\%$ even for the brightest gamma ray sources. Thus, establishing powerful methods of hadronic background rejection is a prerequisite for the effective utilization of observations with the Cherenkov technique.

In this chapter the tree classification method Random Forest (RF) [Bre01a] with modifications for the application in the g/h-separation and the energy estimation is described in detail. A comprehensive study of methods for multi-dimensional event classification (with g/h-separation as chosen application field) showed, that results obtained with RF were among the best ones [BCG+04]. This study involved a variety of simple and more advanced classification approaches, which can be grouped into the following categories:

- Regression tree and PDE (Probability Density Estimator) methods:
  C5.0 and CART (clearly outperformed by RF [BCG+04]), Kernel methods and Nearest Neighbours

- Artificial neural networks: NeuNet [Ern06], NNSU (Neural Network with Switching Units), GMDH (Grouped Method Data Handling), MRS and MLP neural networks. A simple 3-layer feed forward configuration (learning by error back propagation) of the NeuNet package yielded results better or comparable to the more advanced neural network methods, but inferior to the Random Forest

- Direct selection, SVM (Support Vector Machines), LDA (Linear Discriminant Analysis), Composite probabilities (this group can be considered inferior [BCG+04])

The simplicity and good performance mainly of the RF g/h-separation (see section 5.4) led to an implementation and intensive usage of the RF method in the analysis of data taken with the MAGIC telescope. A more classical approach of g/h-separation - in analogy to the so-called *Mean Reduced Scaled Width* (MRSW) and *Mean Reduced Scaled Length* (MRSL) [A+05] - will be taken as 'yard stick' to prove the reliability of the Random Forest method for the analysis presented in this work.

The following sections give a complete description of the RF technique using g/h-separation as illustrative object, since existing mathematical treatments show only a few

concrete aspects, if at all. Results and special remarks related to g/h-separation and energy estimation will be given in sections 5.4 and 5.5 respectively.

5.1   Basics of the Random Forest (RF) method

g/h-separation and energy estimation can be seen as multi-dimensional event classification. Imaging Cherenkov telescopes provide a Cherenkov light image, which is object of a further analysis with the so-called Hillas Parameters [Hil85] and some additional image parameters (see section 4.3). These parameters as well as related observational conditions (e.g. the telescope's zenith angle) are the components of an event vector $\boldsymbol{v}$. Now, starting- and end point of g/h-separation and energy estimation can be depicted as follows

- **g/h-Separation**:

    - *Input*:
        * Vector $\boldsymbol{v}$
        * Classes and related labels (*input for training only*):
          γ (0), hadron (1)

    - *Output*:
      Hadronness $h \in [0, 1]$ assigned to $\boldsymbol{v}$, calculated from a mean class label. h can be regarded as 'hadron probability', i.e. the probability, that the event belongs to class 1. h behaves like a probability, since it is calculated from class populations (frequency definition of probability) and covers the interval $[0, 1]$.

- **Energy estimation**:

    - *Input*:
        * Vector $\boldsymbol{v}$
        * Classes and related labels (*input for training only*):
          E_0 (0), E_1 (1), ..., E_{n-1} (n-1)

    - *Output*:
      Estimated energy $E_{est} \in [E_0, E_{n-1}]$, assigned to $\boldsymbol{v}$, calculated from a mean class label $l \in [0, n-1]$ (which corresponds to a mean energy = estimated energy)

In the case of g/h-separation the classes are gammas and hadrons, whereas in the case of energy estimation the classes correspond to energy bins, which are usually chosen to be logarithmically equidistant (there is also an energy estimation version of RF, which does not rely on class-labels. It will be discussed in section 5.5). The vector $\boldsymbol{v}$ introduced above spans a space, which is in general multi-dimensional.

As in many other classification and regression methods, a Random Forest (i.e. a collection of 'randomized' trees) is created using a training sample, which is suitable for the application. In particular, this sample contains the mentioned class labels, which are used as 'optimization target' for the training process.

Using single trees for classification purposes usually gives mediocre results. The tree is overoptimized on the training sample, and there is only poor generalization (new events will be classified rather badly). To circumvent a difficult to handle pruning process [BFOS83], which is the somewhat more classical solution, the Random Forest algorithm uses an alternative generalization approach: *The main intention is to produce an ensemble of uncorrelated trees, which are combined to form a generalized predictor.* In other words, single trees show a rather bad classification performance, but if one uses a set of independent, or uncorrelated, trees and combines them (e.g. by simply calculating a mean classification), the overoptimization is reduced. 'Growing an ensemble and combining by averaging or voting can turn mediocre predictors into world-class predictors' [Bre99].

Two random elements prior to and within the tree growing process serve to approximate ideally uncorrelated trees. The next two sections deal with the related randomization techniques.

5.1.1 *B*ootstrap *agg*regat*ing* (bagging)

There is usually a single data-sample, which is used for the training procedure. A straightforward solution for obtaining independent trees is to split the training sample into as many non-overlapping subsamples as trees are intended to be grown. But there is usually not enough training data available for this approach. This is especially the case if dealing with air shower data, which are always costly to generate (with respect to time and storage space).

A different way is to produce a bootstrap sample for each tree by sampling n-times with replacement from the original training sample containing n events. This procedure guarantees that the distribution of the events in their image parameters is statistically identical for each bootstrap sample (the probability of selecting an event in the 'sampling with replacement' procedure is constantly $1/n$), while these do not contain the same events. It may (and will) happen that some events are taken more than just one time:

The probability of *not selecting* a certain event is equal to $(1 - 1/n)$, which becomes $(1 - 1/n)^n$ when repeating the selection process n-times. From $lim_{n \to \infty}(1 + x/n)^n = e^x$ one can see, that the probability of *not selecting* an event in the bootstrap procedure becomes $e^{-1} \approx 1/3$. Thus, about 2/3 of all training events are taken for each bootstrap sample.

This means for a certain bootstrap sample, that two thirds of the events are (different) original training events and one third are copies (multiple entries).

5.1.2 Tree growing and *random split selection*

The tree growing begins with the so-called root node, which spans the complete image parameter space and which contains the complete bootstrap sample.

In the following the classification is achieved by splitting (or cutting) each node into two successor nodes using one of the image parameters at a time, with a cut value optimized to separate the sample into its classes (in the case of g/h-separation gammas and hadrons). This corresponds to a successive division of the image parameter space into hypercubes.

In order to measure the classification power (separation ability) of an image parameter and in order to optimize the cut value, the Gini-index is taken (which is just a frequently used measure if dealing with tree classifiers similar to the familiar Q-value used in g/h-separation: $Q = \frac{\varepsilon_\gamma}{\sqrt{\varepsilon_h}}$, ε_γ and ε_h are the gamma and hadron acceptances). The choice of the parameter taken for splitting is randomized (details will follow).

The splitting process stops, if the node size (events per node = hypercube) falls below a limit specified by the user, or there are only events of one class (only gammas or only hadrons) left in the node, which therefore cannot be split further. These terminal nodes can also be called elementary hypercubes since they cover the whole image parameter space without intersections (see figure 5.2).

Each terminal node is assigned a class label l by the therein contained training events ($l = 0$ for gammas and $l = 1$ for hadrons in case of g/h-separation).

For nodes, which still contain a mixture of events of different classes, a mean value is calculated taking into account the class populations N_h of hadrons and N_γ of gammas (the original program [Bre01a] uses a majority vote and does *not* calculate mean values):

$$l_{tree\ i,\ term.\text{-}node\ j} = \frac{N_h}{N_h + N_\gamma}$$

Before going into more details, the application of a trained tree (i.e. the classification process) is briefly described:
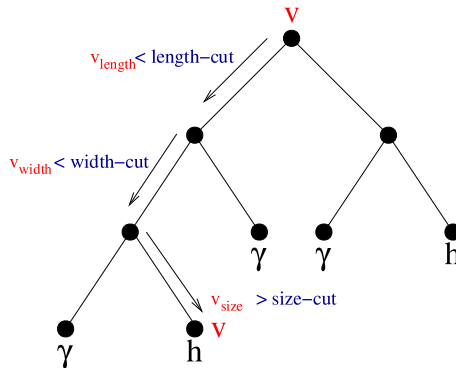
Figure 5.1: Sketch of tree structure showing the classification of an event v by a single tree. The cuts imposed on v at each node are denoted as inequalities. One can see the 'track' of v through the tree which leads to a classification of the event as hadron.

One must take a completely grown tree as starting point (see figure 5.1). The task is to classify a certain event (corresponding to a vector v). Thus, v must be fed into the decision tree: At the first node there is a split in a certain image parameter (for example *length*). Depending on its *length* value the event v proceeds to the left node (*length* < split value) or to the right node (*length* $\geq$ split value). Now the corresponding node again splits in some other (or by chance the same) image parameter. The result is that v follows a track through the tree determined by the numerical values of its components and the tree nodes' cut values, and it will end up in a terminal node.

This terminal node assigns a class label to v (as mentioned above). The label l can now be denoted as $l_i(v)$, where i is the tree number.

After v has been classified by all trees (due to the randomization in general different results will be obtained for different trees, this is the main idea) a mean classification is calculated:

$$h(v) = \frac{\sum_{i=1}^{n_{tree}} l_i(v)}{n_{tree}} \tag{5.1}$$
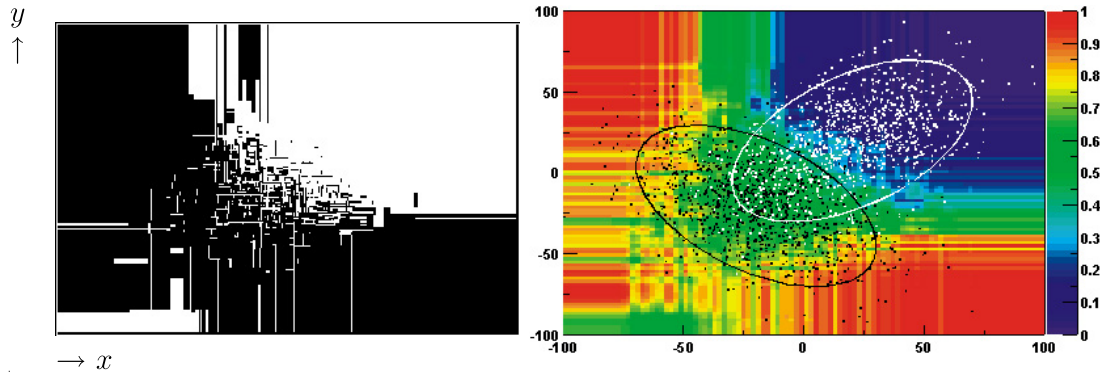
Figure 5.2: Illustration of the Random Forest method: For each of the 2 classes a sample of (x, y) points has been generated according to a 2-dimensional Gaussian distribution, with different mean values $(\bar{x}, \bar{y})$ and covariances $\overline{(x - \bar{x}) \cdot (y - \bar{y})}$ for the 2 classes (see the scattered points and the 1σ contours in the figure on the right). After the training process one obtains a Random Forest 'probability chart' (hadronness). **Left:** A single tree, hadronness can take only values 0 or 1 (for minimum node size = 1). **Right:** A forest of 100 trees, the mean hadronness has been calculated and made visible by means of a color scale.

This mean classification is called hadronness and later on used as the only split-parameter in the g/h-separation (see figure 5.3).

Going back to the training process, the node-splitting proceeds as follows. It is randomized by a feature called random split selection. This means, that parameter candidates for a split are chosen randomly from the total number of available parameters. The parameter and corresponding cut value, which yield the smallest weighted sum of Gini-indices of the 2 successor nodes (see below), is taken for the splitting. The Gini index is a split quality measure named after the Italian economist Corrado Gini.

In the case of two classes the Gini-index Q_{Gini} can be referred to as Binomial variance of the sample scaled to the interval $[0, 1]$. The Gini-index Q_{Gini} can be expressed in terms of the node class populations N_γ and N_h as well as the total node population N:

$$Q_{Gini} = \frac{4}{N} \cdot \sigma_{binomial}^2 = 4 \frac{N_\gamma}{N} \cdot \frac{N_h}{N} = 4 \frac{N_\gamma \cdot (N - N_\gamma)}{N^2} \tag{5.2}$$

Minimization of Q_{Gini} means to minimize the variance of the population of gammas and hadrons which naturally yields a purification of the sample: Q_{Gini} is zero for the ideal
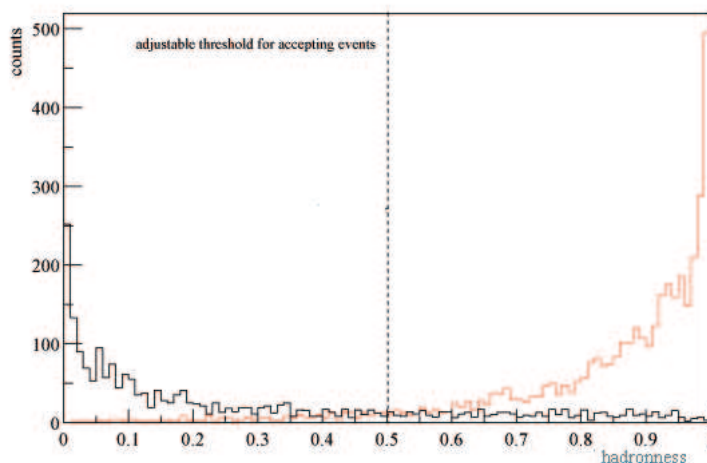
Figure 5.3: Graph showing the mean hadronness for the two test samples of gammas (black curve) and hadrons (red curve). The hadronness is the final cut parameter used in g/h-separation. It is the object for optimization using criteria like the Q-factor.

case that only one class is present in the node ($N_\gamma = 0$ or $N_h = 0$). Figure 5.4 shows the Gini-index as function of the class populations N_γ and N_h.

The Gini-index of the split is calculated by adding the Gini-indices of the two successor nodes (denoted by left and right node) using the node population as weights. Thus, using the abbreviations

$$p_\gamma = \frac{N_\gamma}{N} \tag{5.3}$$

$$p_h = \frac{N_h}{N} \tag{5.4}$$

one obtains (with the indices L and R relating to the left and right node respectively)

$$\overline{Q}_{Gini} = \frac{4}{N_L + N_R}\left(N_L \cdot p_{\gamma L} \cdot p_{hL} + N_R \cdot p_{\gamma R} \cdot p_{hR}\right) \tag{5.5}$$

The minimization of the Gini-index provides the image parameter and the corresponding split value (see figure 5.4, where the Gini-index of one node is drawn). Attempts to improve the arithmetic-mean calculation of the hadronness by using weights were not successful [BCG⁺04]. See e.g. [BL01] for an overview of different split criteria including the Gini index.

5.2   Control of the training process

The following list of items fully covers the steering of the training procedure:
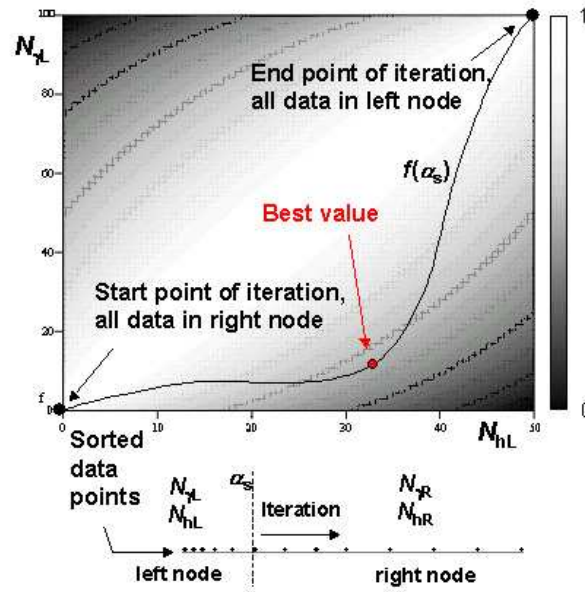
Figure 5.4: The Gini-index Q_{Gini} as function of the class populations N_γ and N_h. As indicated below the plot, an iterative trial of cuts α_s in a parameter α (corresponding to a curve $Q_{Gini}(\alpha_s)$) results in finding the best cut value (with smallest Q_{Gini}).

- *Number of trees*
  The number of trees n_{tree} must be chosen large enough to assure the convergence of the error $\sigma(n_{tree})$, which is the RMS deviation of the estimated hadronness from the true hadronness:

$$\sigma(n_{tree}) = \sqrt{\frac{\sum_{i=1}^{n_{sample}} \left(h_{est}^i(n_{tree}) - h_{true}^i\right)^2}{n_{sample}}} \tag{5.6}$$

Here $h_{est}(n_{tree})$ denotes the estimated hadronness, which depends on the number n_{tree} of combined trees, and h_{true} is the true hadronness of the test sample.

The convergence process is shown in figure 5.5 for a training of RF on a Monte Carlo γ and hadron sample. σ was calculated for an increasing number of combined trees using a test sample, which is disjunct from the training sample.

From figure 5.5 the following practical method can be deduced:

One begins with a reasonably big number of trees, performs the training process and judges afterwards how many trees are to be taken using a test sample. In the case of g/h-separation 100 trees are usually sufficient, if the training samples contain a few 10 thousand events and standard image parameters (*width, length, size, dist*) are taken. The trees generated during the training process are stored successively in a file. Thus, for the classification task one can read in the actually needed number of trees.
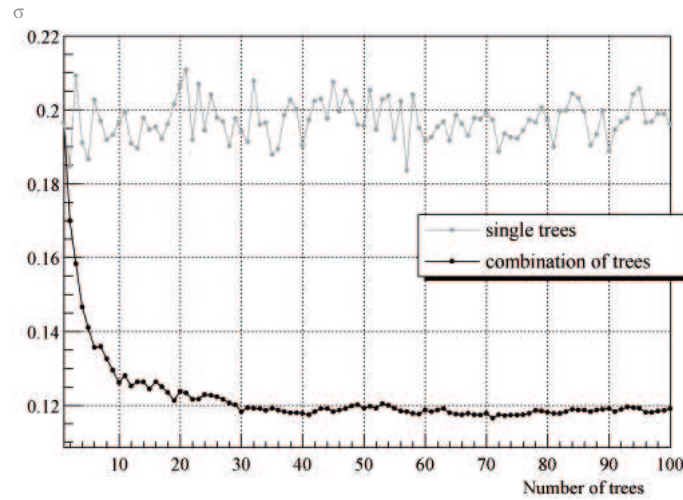
Figure 5.5: The variance σ of estimated hadronness as function of number n_{tree} of combined trees. Also shown is the variance σ for each single tree. A clear improvement due to the combination of trees is visible and a convergence of σ in the case of combined trees is evident.

The above mentioned bagging procedure allows one to judge the σ-convergence also in another way and serves for variable-importance estimates (explanations follow).

When generating a bootstrap sample for a certain tree, about 1/3 of the training sample's events are *not* selected (see introduction of this chapter). These events form the so-called out-of-bag data for the tree (each tree has its own bootstrap sample and therefore its own out-of-bag data), which can be used as test data. The other way around, one can say that each event in the original training sample can be used as test event for about 1/3 of the trees. Thus, if one finds a sufficient convergence of σ (calculated from out-of-bag data) after, say, 300 trees, in fact 100 trees are needed. Since the out-of-bag data are available already in the training stage, they provide immediate σ-estimates.

Some words about *overtraining*:
According to Breiman [Bre02] the overtraining (or overoptimization) vanishes in the case of an infinite number of trees. The practical method described above favours a minimal forest with a number of trees sufficiently large to ensure a classification error (*of the test sample*), which is not significantly decreased by adding more trees.

Such a forest still shows overtraining: When applying g/h-separation *to the training data*, the classes of gammas and hadrons can usually be completely separated by a cut in *hadronness* $= 0.5$. In other words, each tree 'learned the training events by heart' (which is then, to a certain degree, also true for the whole forest). During the tree growing the exact cut values of the parameters are adjusted according to the training sample. This *overtraining* is not a drawback, it affects (of course) merely

the training sample, which provided these exact cut values. The situation can be compared to adjusting cuts in image parameters like the Scaled Hillas Parameters (see section 5.4): One optimizes the cut values on a certain gamma source (i.e. using real, observational data) or on Monte Carlo data and later on applies them to the data, one actually wants to analyze (and which *must be completely disjunct* with the training data).

As a matter of course, the performance of RF or any other g/h-separation method must be checked using a test sample.

- *Number of trials in random split selection*
  The second element, which (besides the bagging procedure) introduces a certain degree of randomization to the training process, is the random split selection. The number of selected image parameters, which serve as candidates for a split, will be called in the following 'number of trials' N_{try}.

  A good empirical value is $N_{try} = \sqrt{N_{par}}$ where N_{par} is the total number of parameters used in tree growing [Bre02].

- *Node size = minimum size of node at which further splitting stops*
  For correctly labeled training events node size = 1 is preferable [Bre02], for partly incorrect labeled data (e.g. using ON-data as hadrons) it is advisable to take a node size > 1, since data are not intended to be split completely. Experience tells that still a small number < 10 is best.

5.3 Basic mathematical description of the Random Forest Method

For a detailed mathematical treatment of the Random Forest method the reader is referred to [Bre01b]. Here, a short introduction to the relevant mathematical terms used when dealing with predictor ensembles as well as the main conclusions are given.

The output of a single tree grown according to a random vector θ is denoted as $h(\theta, \boldsymbol{v})$. $\boldsymbol{v}$ is the above mentioned input vector. The vector θ is a 'map' of the tree structure, i.e. it contains a ordered map of the nodes together with the image parameters and split values related to each node. l is the true class label of the event $\boldsymbol{v}$. It is common to introduce the following definitions [Bre01b]:

- **Margin**
  The so-called margin is defined as:

$$mr(\boldsymbol{v}, l) = \frac{1}{n_{tree}} \sum_{k=1}^{n_{tree}} I(h(\theta_k, \boldsymbol{v}) = l) - \frac{1}{n_{tree}} \max_{j \neq l} \left(\sum_{k=1}^{n_{tree}} I(h(\theta_k, \boldsymbol{v}) = j) \right) \quad (5.7)$$

Here I denotes the identity function, i.e.

$$I(h = l) = \begin{cases} 1 \text{ if argument is true, i.e. } h = l \\ 0 \text{ if argument is false, i.e. } h \neq l \end{cases} \quad (5.8)$$

For a 2-class problem the margin can be written as

$$mr(\boldsymbol{v}, l) = \frac{2}{n_{tree}} \sum_{k=1}^{n_{tree}} I(h(\theta_k, \boldsymbol{v}) = l) - 1 \qquad (5.9)$$

Thus, the margin can be understood as the proportion of correctly classifying trees mapped to the interval $[-1, 1]$ instead of $[0, 1]$.

• **Strength**
The expectation value of the margin is called strength.

$$s = E_{\boldsymbol{v}, l}(mr(\boldsymbol{v}, l)) \qquad (5.10)$$

The indices $\boldsymbol{v}$ and l indicate a calculation of the expectation value summing over a test sample containing events and the corresponding labels $(\boldsymbol{v}, l)$. In the following also the variance σ^2 and probabilities P are calculated in this way.

• **Generalization Error**
$$PE = P_{\boldsymbol{v}, l}(mr(\boldsymbol{v}, l) < 0) \qquad (5.11)$$

The generalization error PE can be understood as classification error if a *plurality vote* is used, which is in analogy to a cut in hadronness at 0.5. PE is converging to a limiting value, which follows from the law of large numbers (applicable due to randomization of tree growing). This means that a possible overtraining will vanish, when more and more trees are added to the forest.

An upper bound of the generalization error PE can be derived from the Chebychev inequality. One obtains [Bre01b]:

$$PE \leq \frac{\sigma_{\boldsymbol{v}, l}^2(mr(\boldsymbol{v}, l))}{s^2} \qquad (5.12)$$

$$PE \leq \overline{\rho} \cdot \frac{1 - s^2}{s^2} \qquad (5.13)$$

Here $\overline{\rho}$ denotes the mean correlation between the trees and s is the strength as introduced above. The mathematical implication can be formulated in the following way: *The generalization error (with its interpretation as misclassification rate) is kept small by a low correlation of the trees (high degree of randomization) and a big strength.* A big strength is given by good predictor properties of each single tree, which is mainly determined by powerful image parameters, i.e. image parameters which show different (non-overlapping) distributions for gammas and hadrons.

The following two sections deal with the application of RF in g/h-separation and energy estimation. The related studies (mainly concerning the control of the training process and the choice of the image parameters) were carried out using pure Monte Carlo data in order to exclude systematic errors, which could arise from not perfectly matching MC and observational data. Nevertheless, the analysis methods (including RF) will be tested on data taken during observations of the Crab Nebula - a steady gamma source, which serves as the 'standard candle' in gamma ray astronomy (see chapter 6).

5.4 Application of RF in g/h-separation

The last sections addressed basics of the RF method, referring to g/h-separation as illustrative object. Now, some more particular features will be discussed (which are partly valid also for other g/h-separation techniques where a training process is involved), and a comparison with a standard g/h-separation approach is carried out. Before discussing these studies, some clarifying explanations and definitions are given.

Training sample

The gamma training sample is obtained by means of Monte Carlo simulations. However, the hadron training sample may be composed of Monte Carlo, OFF or ON data.

It is usually not advisable to use MC hadrons when analyzing observational data, since hadronic showers are more complex than gamma showers (i.e. MC simulations may be more error-prone) and typically subject to very large fluctuations. The low trigger probability makes hadronic simulations additionally time consuming. On the other hand gamma showers have a pure electromagnetic nature, which is well understood and comparably easy to simulate. In fact there is no need to use MC hadrons, when OFF or ON data is available. The small fraction of gammas, which is contained in ON data, does not hinder to use it as hadron sample. This is implied by the fact, that Random Forest is stable enough to deal with an impure hadron sample containing up to 5 %-10 % of gammas (see figure 5.6). Even higher 'degrees of impurity' do not deteriorate the performance for strong g/h-separations (i.e. low hadron acceptances). The related study was carried out by successively contaminating a proton MC sample with MC gammas (details concerning the MC samples are given below). The RF training and test was done using the following image parameters:

- $\log_{10}(size)$, *dist*
- *width*, *length*
- *dens*
- $\cos\theta$ with θ being the zenith angle
- advanced image parameters: $sign_q$, *relative island size* and *dens*.

If different image parameters are used in the training of the RF, this study must actually be redone. But one can imagine that the most efficient parameters (which are different in their distribution for gammas and hadrons) are already included in the list above, so that new parameters or slightly modified ones should not lead to significantly different results.

Thus, when dealing with observational data, one can use ON data as the hadron sample. However, as already mentioned, one main concern of this chapter is the comparison of the RF with a standard technique in g/h-separation. Thus the studies were done using merely MC data, in order not to introduce systematic effects due to not perfectly matching MC and observational data.
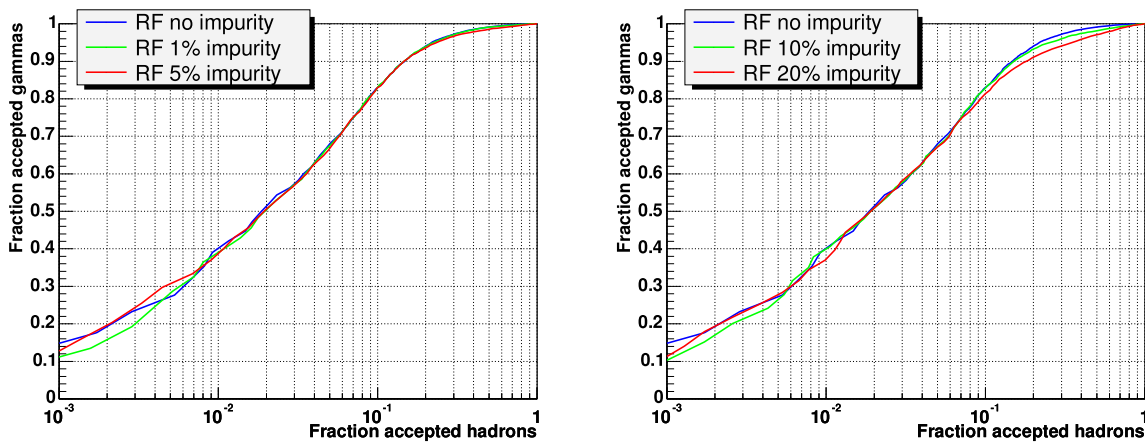
Figure 5.6: Neyman-Pearson plots showing the effect of partly wrongly labeled hadron data. It is clearly visible that the hadron sample containing 5 % gammas (randomly selected from a Crab Nebula like energy distribution, see text) has only a small influence on the performance of the RF.

Types of image parameters

All parameters are treated in the same way, which means that in particular detector related or observational parameters like $\cos\theta$ (θ is the zenith angle) must be used with care. The intention of using such parameters is, that cuts in other image parameters depend on them. It shall be mentioned, that also the *size* parameter (integrated signal of the image) may play this role. Thus, in a more general scheme, one can distinguish between parameters, in which one wants to cut, and parameters, on which the cuts may depend.

Therefore, the training data should be chosen not to permit a classification using these parameters alone, e.g. by using *the same distribution* of $\cos\theta$ in the training samples of both classes. However, one can have differently effective g/h-separations for different zenith angle intervals. When cutting in hadronness, a certain range of zenith angles may be preferably selected. However, this was not observed for the analysis presented in this work, where the zenith angle range is restricted to $0° < \theta < 30°$.

Additional attention must be paid if e.g. only one of the samples shows discrete $\cos\theta$ values for technical reasons (limitation in Monte Carlo production). In this case the $\cos\theta$ values must be rounded exactly in the same way also for the other sample.

The so-called *alpha* parameter (see sections 2.2, 4.3 and 4.5 as well as [BCG⁺04]), which describes the image orientation in the camera, is the strongest g/h-discriminator, since gamma images point to the camera source position whereas hadronic images show a random orientation: The shower axes of hadronic cosmic ray showers are isotropically distributed, which is reflected in the image orientation. The *alpha*

parameter is *not* considered in this study, since it plays a particular role in proving the existence of a signal and determining its strength (see section 4.5).

The following MC data set was taken for further investigations of RF:

- *Proton simulation settings*
  Energy range: $200\,\text{GeV} < E < 30\,\text{TeV}$
  Zenith angle range: $0° < \theta < 30°$
  Spectral index: $\alpha = -2.7$
  Impact parameter: $R < 400\,\text{m}$
  Viewcone: $\delta < 5°$

- *Gamma simulation settings*
  Energy range: $50\,\text{GeV} < E < 30\,\text{TeV}$
  Zenith angle range: $0° < \theta < 30°$
  Spectral index: $\alpha = -2.6$ (Crab Nebula-like energy spectrum)
  Impact parameter: $R < 200\,\text{m}$

Due to limitations in the MC production of protons (which are very time consuming to produce due to the low trigger rate), a MC production with higher energy threshold than the standard one was taken. There is no drawback, since this chapter is merely concerned with a comparison of the separation using the RF method and the separation based on Cuts in Scaled Hillas Parameters, which are an established g/h-separation method for the higher energy range starting at about $100\,\text{GeV}$.

The MC data were treated with the same analysis methods (in particular image restoration and image parameter calculation, see last chapter) as real, observational data. The figures 5.8 and 5.9 show the *width* and *length* distributions of the MC data introduced above after application of cuts in *leakage* and *dist*. The sudden change in the slope of the MC proton profile plot is caused by a high muon content in the *size* region right below $\sim 1000\,\text{phe}$. These muons distort the *width* and the *length* Hillas Parameters towards large values.

The cut in *leakage* < 0.1 removes events, which show a concentration of more than $10\,\%$ of the image size in the outermost pixel ring. If not rejected, such events would - apart from a general distortion of the image parameters - introduce an error of possibly more than $10\,\%$ in the energy estimation only due to the truncated *size* and distorted *dist*. An additional cut in *dist* $> 0.3°$ removes nearly-round images, which show a badly resolvable orientation, i.e. the values for the *alpha*-parameter are widely spread (see figure 5.7).

The strong dependencies of the *width* and the *length* parameters on the *size* or $\log_{10}(size)$ of the image are evident. Therefore, the most simple method of g/h-separation, which is the direct use of cuts in the image parameters (with the *width* and *length* of the Cherenkov light image being the most effective parameters), must take into account mainly this *size* dependence. Thus, differently adjusted cuts must be used for each *size* bin (so-called dynamical cuts). As an alternative approach Scaled Hillas Parameters in analogy to the so-called Mean Reduced Scaled Hillas Parameters [A$^+$05] are introduced, where the *size* dependence is 'scaled out':

MC protons   *MC gammas*
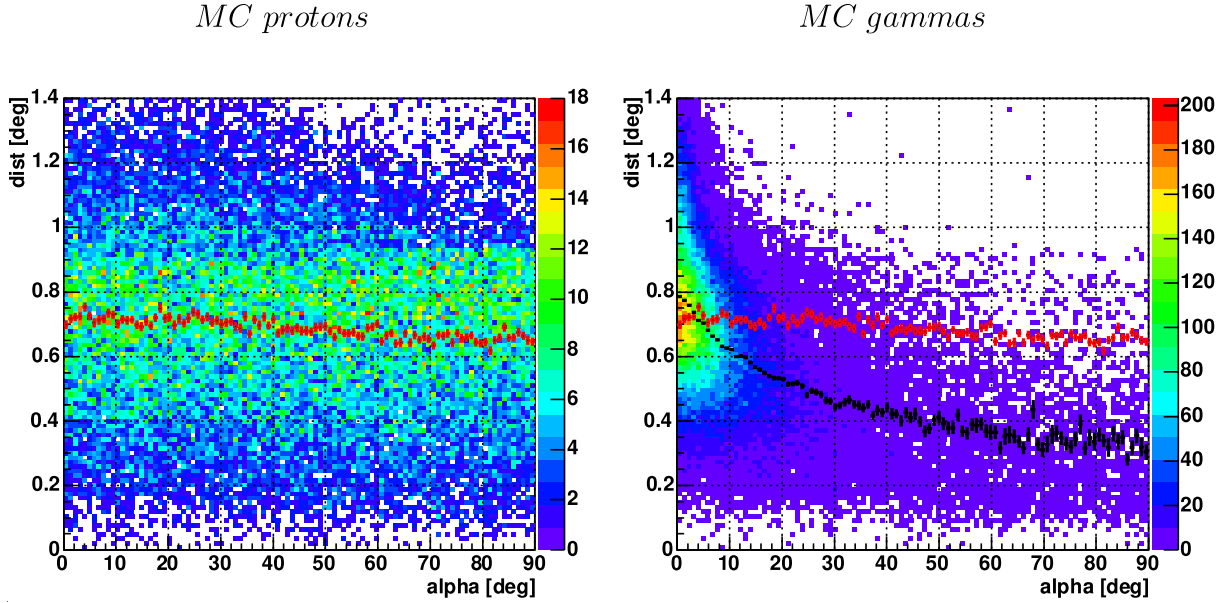
Figure 5.7: *dist* versus *alpha* distributions for MC protons (on the left) and gammas (on the right). Additionally, the profile plots of gammas (black) and protons (red) are shown. A cut of $dist > 0.3°$ removes events with wide-spread *alpha* values.

One divides the *size* range into bins. For each bin i the mean $\overline{w_i}$ and the variance σ_{w_i} of the *width* distribution are calculated from the gamma MC. Then the scaled width $w_{i,scaled}$ of each bin is obtained by:

$$w_{i,scaled} = \frac{w_i - \overline{w_i}}{\sigma_{w_i}}$$

The same procedure is imposed on the *length* parameter. As a result one obtains rectified *width* and *length* distribution for gammas or, to be more precise, these distributions are standardized, i.e. $w_{i,scaled}$ and $l_{i,scaled}$ follow a probability density function (pdf) with mean 0 and variance 1. This procedure was carried out independently for the three different zenith angle bins $0° < \theta < 10°$, $10° < \theta < 20°$ and $20° < \theta < 30°$. The resulting *mean reduced scaled width* and *length* parameters are shown in figures 5.10 and 5.11 (compare with figures 5.8 and 5.9). This kind of scaled parameters will be hereafter called *Scaled Hillas*. As a result one can use in the g/h-separation *size*-independent static cuts in the *scaled width* and *length*.

The cut optimization and the training of the RF demand training samples, which should principally not differ from real observation conditions. This means in particular, that the zenith angle distributions of gammas and protons must match. In order to remove a possible dependence of the g/h-separation on the gamma-spectrum the *size* distributions of gammas and protons should match as well.

Taking this into account, training samples containing each 15000 events for gammas and hadrons were randomly selected (the zenith angle and *size* distributions of the original

sample already showed similar θ- and *size* distributions). Special care was taken to use values of $\cos\theta$ rounded to 2 decimal digits for both gammas and protons (originally only the proton sample showed this rounding).

In the case of RF the training process was described in the previous sections. Also the cuts in Scaled Hillas Parameters must be adjusted in a training or optimization process. Only upper cuts in *scaled width* and *scaled length* were evaluated, which is completely sufficient in the region of *size* > 200 phe, where the length and width distributions show only a small overlap. As optimization function the Q-value was chosen.

$$Q \;=\; \frac{\varepsilon_\gamma}{\sqrt{\varepsilon_h}} \tag{5.14}$$

$$\varepsilon_\gamma \;=\; \frac{N_{\gamma,sel}}{N_\gamma} \tag{5.15}$$

$$\varepsilon_h \;=\; \frac{N_{h,sel}}{N_h} \tag{5.16}$$

Here $N_{\gamma,sel}$ denotes the number of selected gammas, which remain after the cut is applied, and $N_{h,sel}$ is the number of remaining hadrons. Thus a maximization of Q as a function of cuts in *scaled width* and *scaled length* yields a high gamma acceptance ε_γ and low hadron acceptance ε_h.

When using real data the significance S is favoured as optimization function. However, for a small number of excess events N_{ex} (i.e. for a weak source) the significance behaves like the Q-value (see section 4.5 and chapter 6, equation 6.1). Furthermore, this chapter is dedicated to a pure MC study and therefore the Q-value is chosen.

The variation of cut values to maximize Q (or, equivalently, to minimize $-Q$) was performed using the minimization package TMinuit, which is part of the ROOT analysis environment [BR06]. It turned out, that the standard minimization - the MIGRAD package of TMinuit, which relies on gradient values to find the steepest descent - failed to find the global minimum. The source of this problem is: *Q is not a continuous function of the cut values, because the event samples change, if the cut values are changed.*

Thus, instead of using the standard MIGRAD algorithm, a Monte Carlo search (the so-called Metropolis package of TMinuit) followed by a simplex minimization (the SIM-PLEX package of TMinuit) proved to provide stable results. The simplex method is slower than any gradient method like MIGRAD, but more robust since it does not rely on functional derivatives and can operate when MIGRAD fails, see [Bra92] for details on different minimization algorithms. One may also think of using more advanced evolutionary strategies like simulated annealing, but the procedure described above is simple and not very time consuming in the case, that only two parameters (upper cuts in *scaled width* and *scaled length*) have to be optimized.

Both the Random Forest and the scaled Hillas method were trained and applied to test data disjunct from the training samples. In the next section the results are compared by means of so-called Neyman-Pearson diagrams (also called 'Receiver operator characteristic' (ROC) curves [BCG$^+$04]), which show the gamma acceptance as a function of the hadron acceptance.

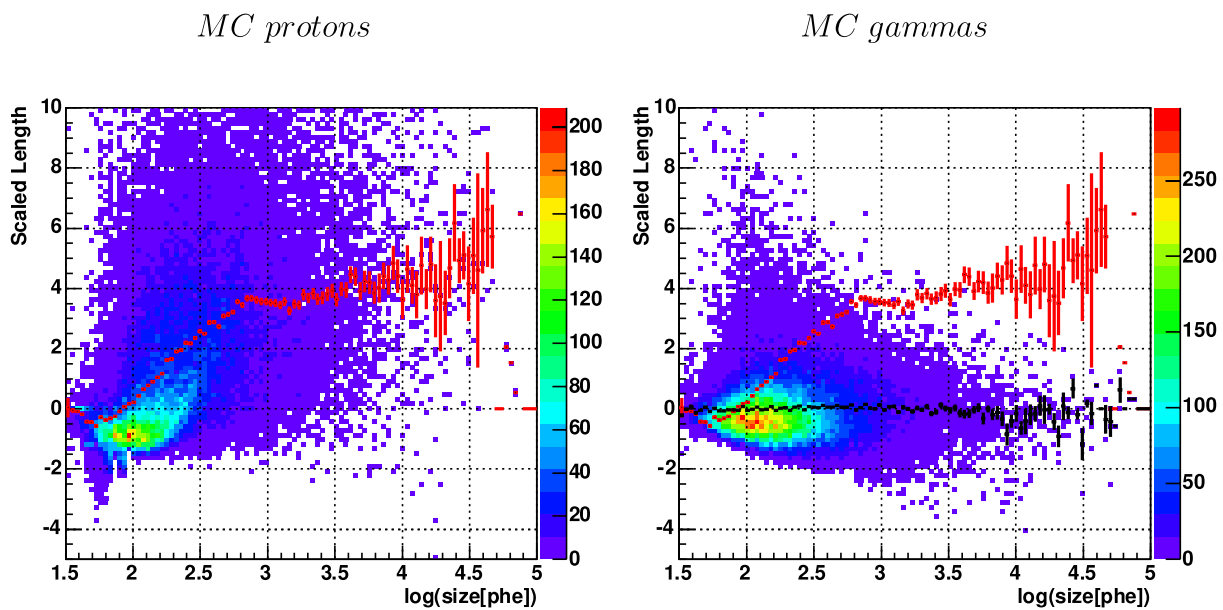*MC protons*   *MC gammas*
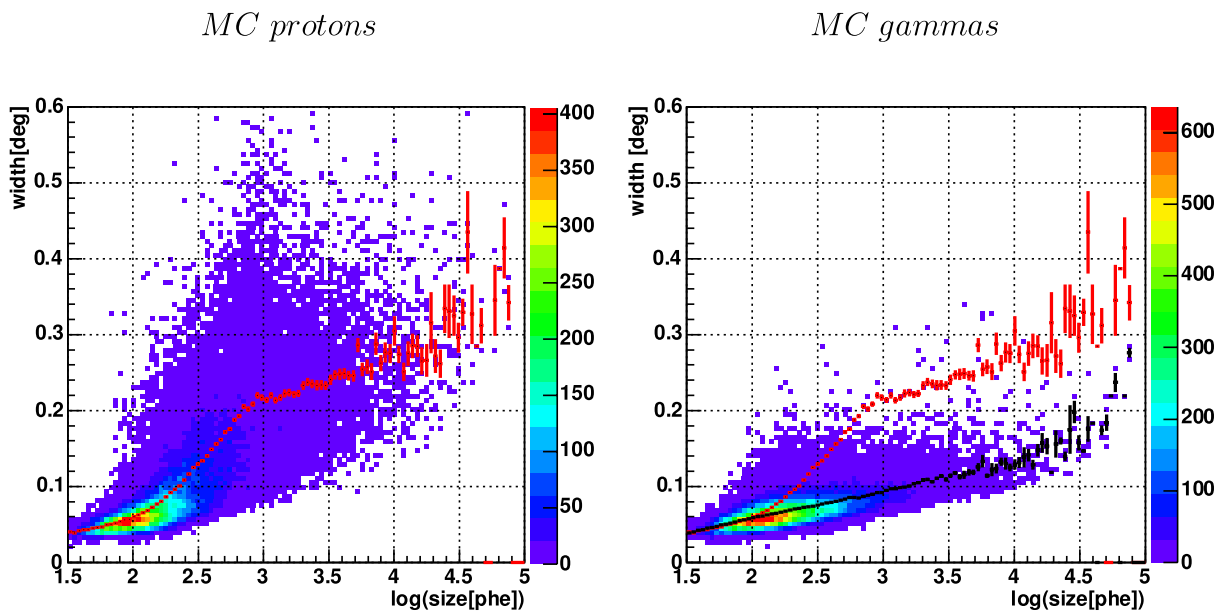
Figure 5.8: *width* distributions of MC protons (on the left) and gammas (on the right). The profile plots (red for protons and black for gammas) show the *width* mean value for $\log_{10}(size)$ bins as indicated by the horizontal error bars (nearly identical with the line width and therefore hardly visible). The vertical error bars are identical with the RMS of the width distribution in the related $\log_{10}(size)$ bin. For comparison, the profile plot of MC protons is overlaid on the gamma width-$\log_{10}(size)$ histogram.

MC protons   *MC gammas*
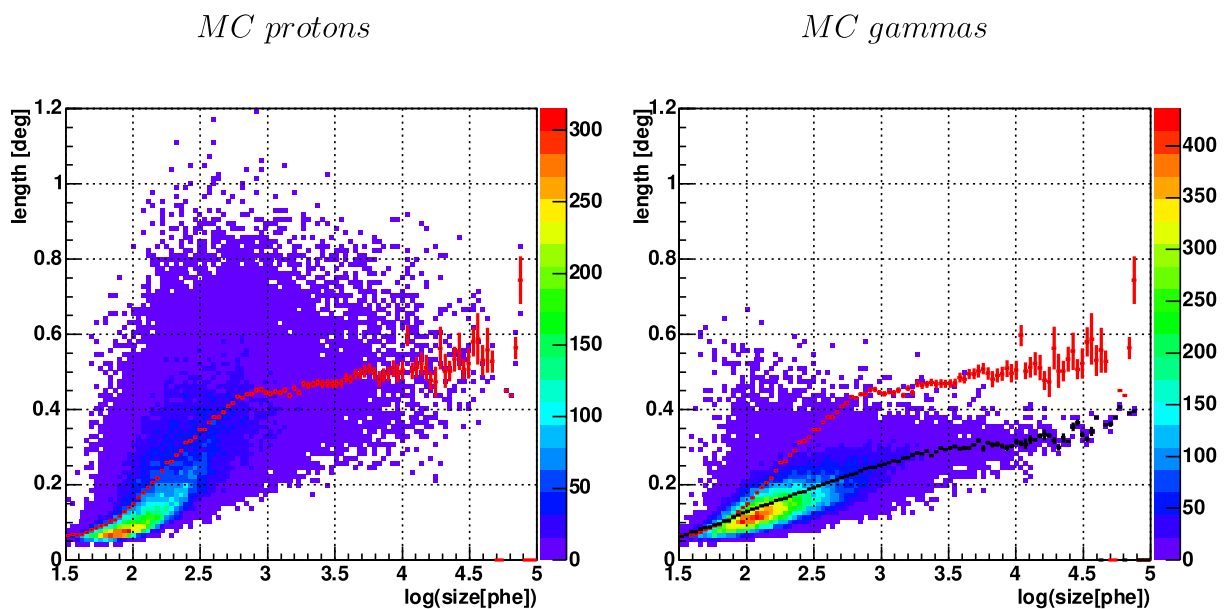
Figure 5.9: *length* distributions of MC protons (on the left) and gammas (on the right). For more explanations see figure 5.8.
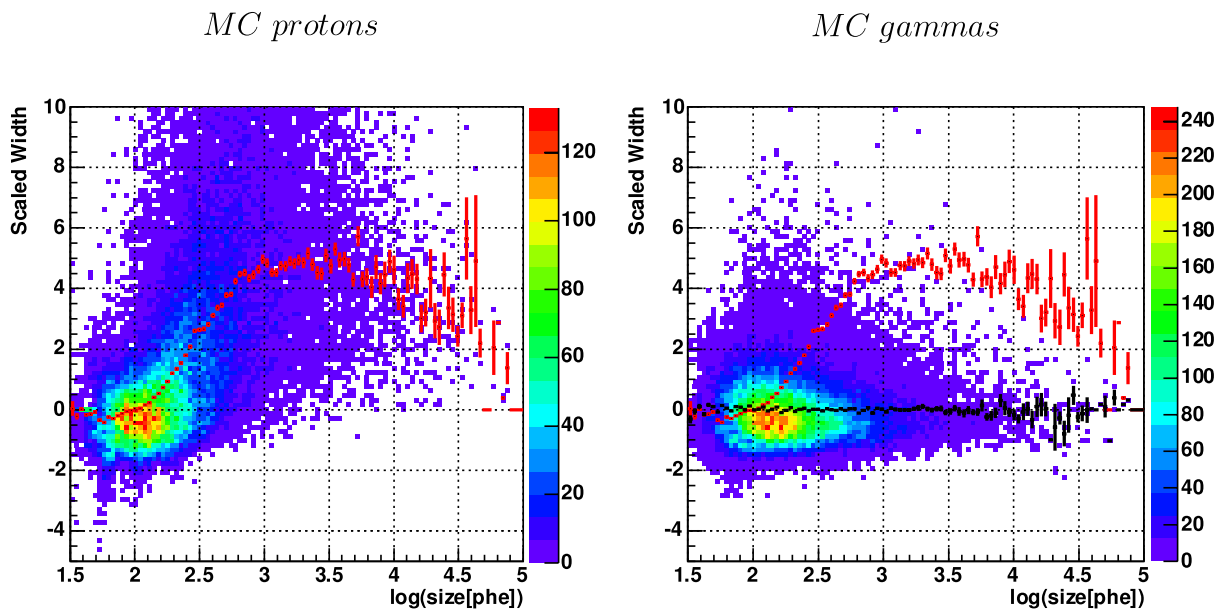
Figure 5.10: *Scaled Width* distributions of MC protons (on the left) and gammas (on the right). For more explanations see figure 5.8.

Figure 5.11: *Scaled Length* distributions of MC protons (on the left) and gammas (on the right). For more explanations see figure 5.8.

5.4.1   Performance of the RF g/h-separation

Unless stated otherwise, in agreement with the theoretical discussion of the last sections, the following settings were used for the RF training:

- Node size = 5

- Number of trials in random split selection = $\sqrt{n_{par}}$ with n_{par} being the number of parameters

- Number of trees = 100

- Size of the training sample = 30000
  (i.e. 15000 gammas and 15000 protons, limited by proton MC sample size)

If not explicitly stated otherwise, most of the results are presented for a threshold in *size* of 200 phe. The source analysis presented in this work (see chapters 6 and 7) showed that there is only a marginal excess below 200 phe. Furthermore, a comparison with the standard g/h-separation approach of cuts in Scaled Hillas is simplified when restricting *size* to values exceeding 200 phe, since then upper cuts in the Scaled Hillas are sufficient. As discussed in the last section, quality cuts in *leakage* and *dist* were applied.

Figure 5.12 shows ROC curves for a RF training on $\log_{10}(size)$, *dist*, *width*, *length* and $\cos\theta$ (std. parameter) as well as the ROC curves for a training on all image parameters as addressed later in this section. It is clear, that the performance of RF is fully exploited when using a sample size of 30000 events.

The next figure documents the stability of the RF performance and its independence on the RF initialization (in particular the initialization of the random number generator). Thus, despite that RF is not a analytical method it provides stable results - a strength derived from the combination of several randomized trees.

In principle, trees are able to deal with correlated input parameters, so that RF should be able to cope with the strong dependences between *width* and *size* as well as *length* and *size*. However, it is known that trees perform worse if curved boundary regions must be approximated by rectangular partitions (c.f. figure 5.2). Thus, it is advisable to test the *scaled width* and *length* parameters as input for RF. As one can see from figure 5.14 it is sufficient if the standard (unscaled) image parameters are used in the RF training.

As introduced in section 5.4 there are different kinds of image parameters: Parameters, which are meant for cutting and those, where the cuts may depend on. First of all, as already mentioned several times, most image parameters and related cut values are strongly *size*-dependent. It may therefore be advantageous to train RF for different *size* regions separately instead of taking *size* as input parameter for RF (with equalized distributions for gammas and hadrons). In the figures 5.16 one can see results of a RF training and test for the *size* regions $200 - 500$ phe and for the *size* region larger than 500 phe. For comparison the ROC curve for an RF training on the whole *size* range is additionally given in both figures. As expected, if *size* is given as RF input parameter with identical *size* distributions for gammas and hadrons, a dedicated training for different *size* bins is not necessary. The same behaviour holds for the zenith angle as input parameter, although here just a weak dependence can be seen, since the MC sample
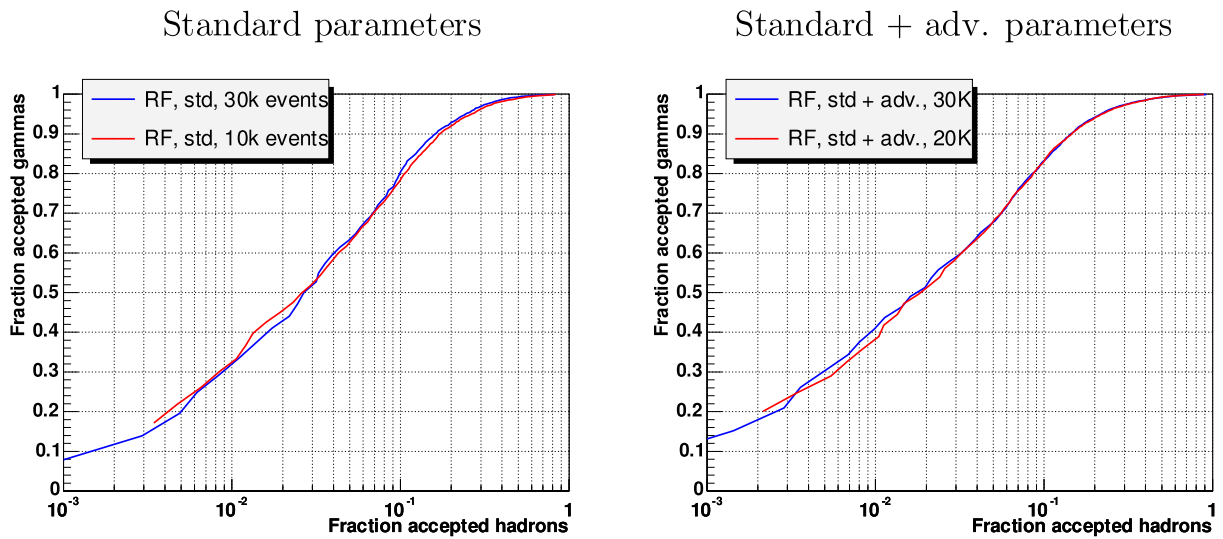
Figure 5.12: Dependence of the RF performance on the training sample size. The horizontal axis (hadron acceptance) is plotted logarithmically for a better visibility of the low hadron-acceptance region. **Left:** The two curves correspond to a number of training events of 10000 and 30000 events (equally split among gammas and protons) as indicated in the legend. The training was done using standard parameters (see text). **Right:** The two curves correspond to a number of training events of 20000 and 30000 events. Here, the RF was trained on standard plus all advanced image parameters discussed in this section. The agreement of the ROC curves indicates that the training sample size of 30000 events used in the ongoing studies is sufficient and does not limit the RF performance.
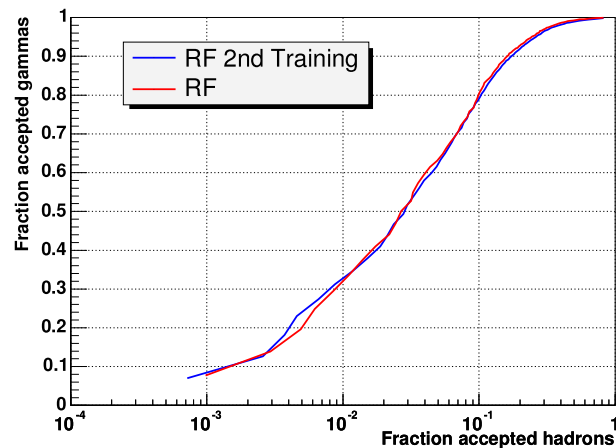
Figure 5.13: Independence of the RF performance on the random number initialization. The two curves correspond to different seed values for the RF training process.
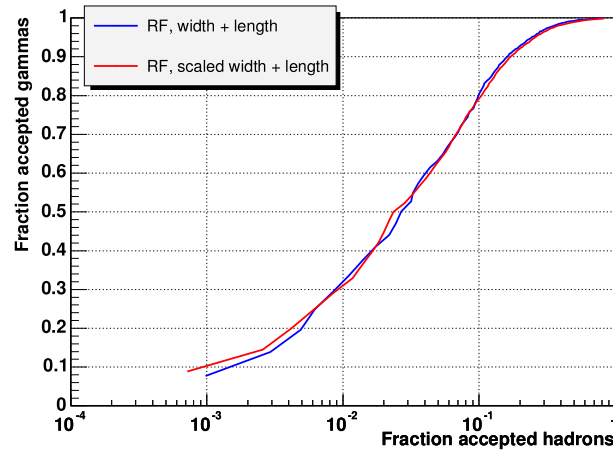
Figure 5.14: ROC curves for RF g/h-separation with scaled Hillas and unscaled *width*, *length* as input. As one can see, RF can handle the *width* and *length size*-dependence without need of the scaled Hillas Parameters as input.
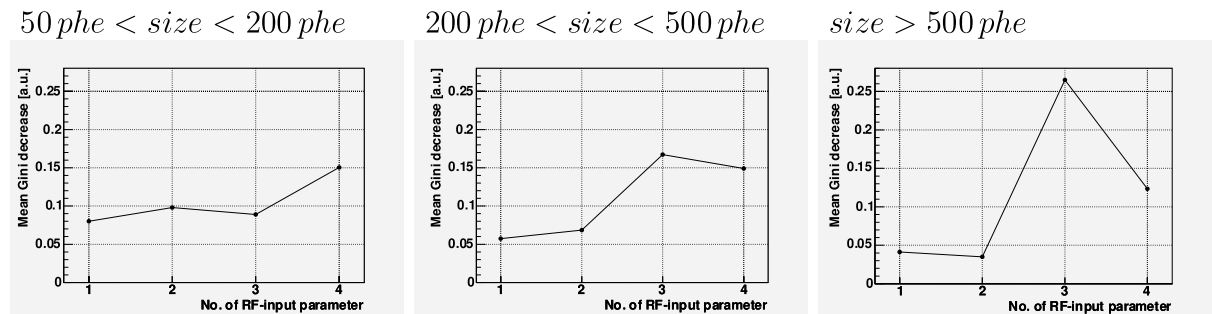
Figure 5.15: Variable importance graphs for RF separately trained in different *size* regions. From left to right: *size* range $50 - 200$, $200 - 500$ and larger than 500 phe. The x-axis shows the parameter index: $1 \stackrel{\triangle}{=} \boldsymbol{size}$ $2 \stackrel{\triangle}{=} \boldsymbol{dist}$ $3 \stackrel{\triangle}{=} \boldsymbol{width}$ $4 \stackrel{\triangle}{=} \boldsymbol{length}$. One can see, that the *width* and *length* parameters change their importance (separation power) when going from the low *size* region of $50 - 200$ phe to the higher *size* range of *size* > 200 phe. This fact is in agreement with figures 5.9 and 5.8, where the width distribution of gammas and protons merge already at higher *size* values than the respective *length* distributions.
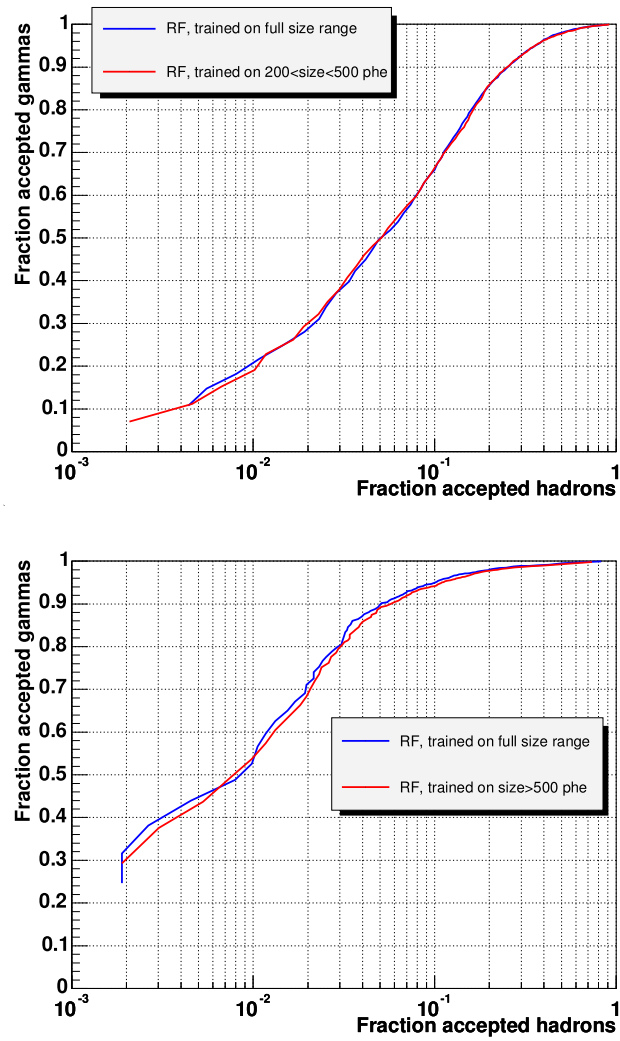
Figure 5.16: ROC curves showing the performance of RF if explicitly trained on the given *size* ranges. The comparison with a random forest trained on the whole *size* range shows comparable results.
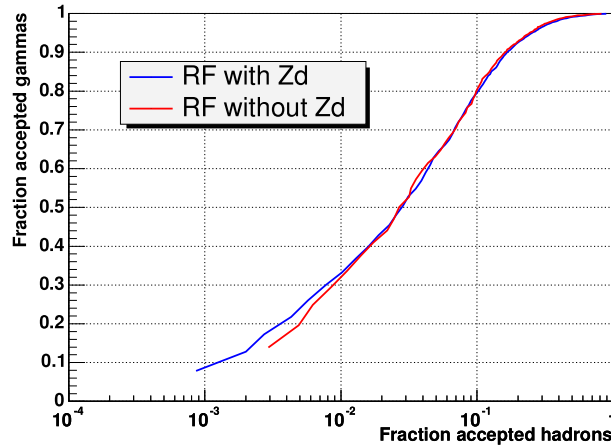
Figure 5.17: Comparison of RF results with and without introducing $\cos\theta$ as input parameter to RF.

under investigation covers merely the zenith angle range from 0 to 30 degrees (see figure 5.17). The figures 5.15 show the 'mean Gini decrease' with $\Delta Q_{Gini} = Q_{Gini,\,before\,split} - \overline{Q}_{Gini,\,after\,split}$ for the different RF input parameters in different *size* bins. The mean of ΔQ_{Gini} is calculated by averaging over all nodes of all trees, where the corresponding parameter is taken for cutting. A high value of ΔQ_{Gini} for a certain parameter indicates, that this parameter has a high separation power (it serves for a decrease in Gini-index, i.e. a decrease of the binomial sample variance). Thus, a diagram showing ΔQ_{Gini} as a function of the parameter index is called variable importance graph [Bre02].

The following plots show results from the RF g/h-separation when introducing various new image parameters (see section 4.3 for details):

- *dens*

- *relative island size* (islands are spot like image parts completely separated from the main island, a typical hadronic signature).

- *sign$_q$* (derived from the charge asymmetry, which is based on a vector from the image's center of gravity to the charge-square-weighted center of gravity, see section 4.3.1): The intention is that events are accepted, when they show a gamma-like charge asymmetry, i.e. a pointing towards the camera center (the head-tail information of the charge asymmetry is used).

As can be seen from figure 5.18, the new set of parameters introduces a significant improvement in all *size* regions. For a certain hadron acceptance a gain in the gamma acceptance for low and (most notable) high *size* values is clearly visible (see plots on the right in figure 5.18, a gain of, for example, 20 % at a certain hadron acceptance means that the related gamma acceptance is increased by a factor 1.2)):
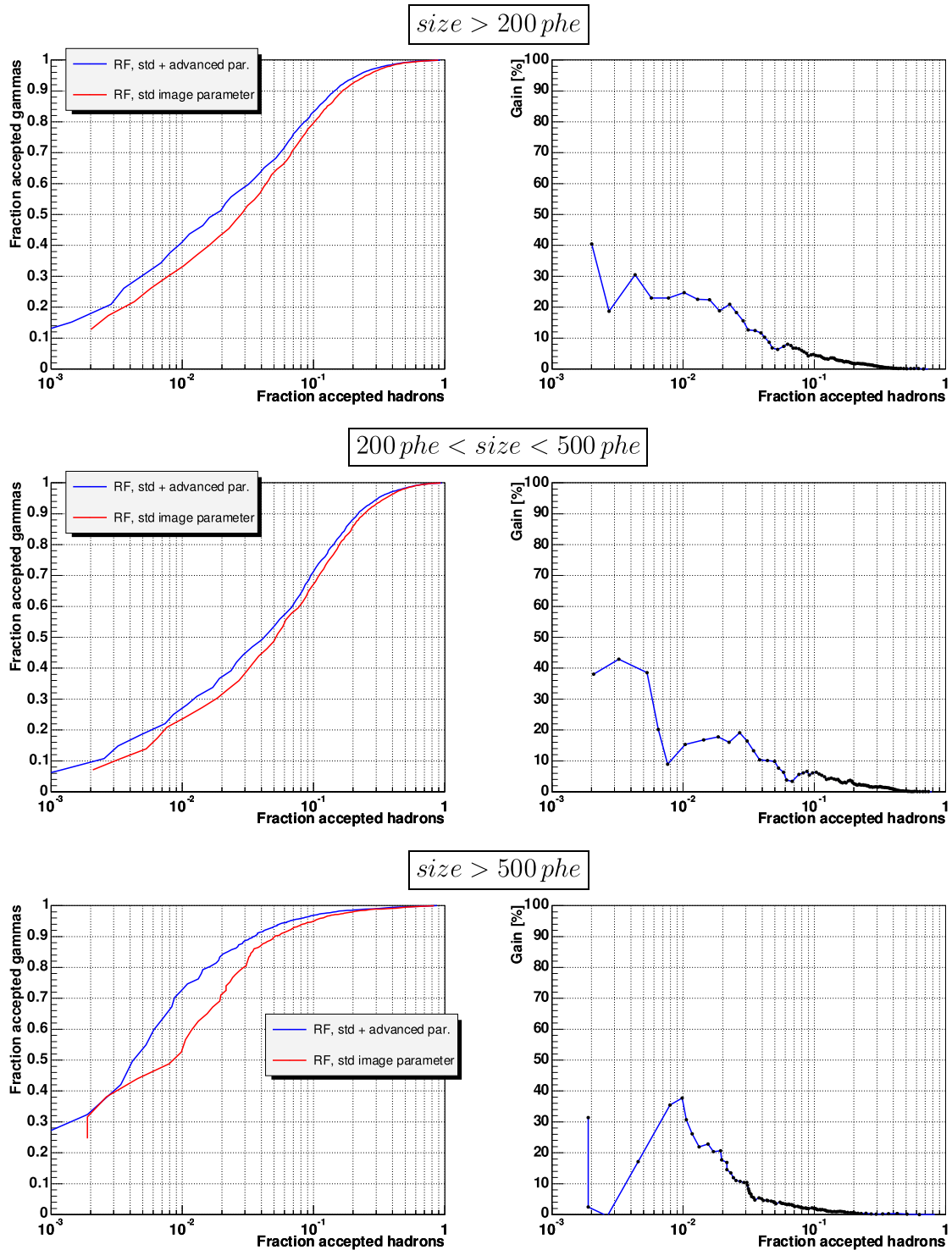
Figure 5.18: Performance of RF for a training on the following parameters: 'Standard': $\log_{10}(size)$, $dist$, $width$, $length$, $cos\theta$. 'Add. advanced par.': $sign_q$, $relative\ island\ size$ and $dens$. The ROC curves correspond to different $size$ regions as indicated. The plots on the right show the gain in the gamma acceptance - introduced by the new set of image parameters - as function of the hadron acceptance.

- 200 phe $< size <$ 500 phe: A gain in the gamma acceptance of about 15 % with maximum values of 40 % is achieved.

- $size >$ 500 phe: A gain in the gamma acceptance of about 20 % with maximum values of up to 35 % is achieved (for hadron acceptances below 0.01 no conclusions can be drawn due to poor statistics).

5.4.2   Comparison between RF and a classical g/h-separation approach

This section deals with the comparison of RF and the g/h-separation method of cuts in Scaled Hillas Parameters.

In a first study Random Forest was trained using the image parameters $\log_{10}(size)$, *width* and *length*. *dist* is neglected as input parameter in order to make a 'fair' comparison (*dist* was not entering into the calculation of the Scaled Hillas due to limited MC statistics). The results of the RF training and test were compared to the cuts in the Scaled Hillas Parameters optimized with the above described procedure.

In order to obtain more than just one point in the ROC diagram, the minimization of $-Q$ was modified by adding the following regularizer R

$$R_{\varepsilon_h}(p) = \left(\frac{\varepsilon_h - p}{\varepsilon_h}\right)^2 \qquad (5.17)$$

so that the minimization condition reads

$$-\left(Q + a \cdot R_{\varepsilon_h}(p)\right) = min \qquad (5.18)$$

Here, a and ε_h are constants. In particular, ε_h must be set to the desired hadron acceptance. p denotes the hadron acceptance as calculated for the current Scaled Hillas cut set, which varies during the minimization process. For typical Q-values of order ~ 10 a high number of $a = 1000$ forces the minimization algorithm to surely put the regularizer to 0 (by adjusting p to ε_h, which is the actual aim). In a kind of 'fine tuning' the cut values in *scaled width* and *scaled length* are furthermore chosen to obtain a maximized Q (always keeping $p = \varepsilon_h$). Figure 5.19 shows the results. Improvements of more than 20 % or even 40 % (RF training including *dist*) over a wide range of hadron acceptances are clearly visible.
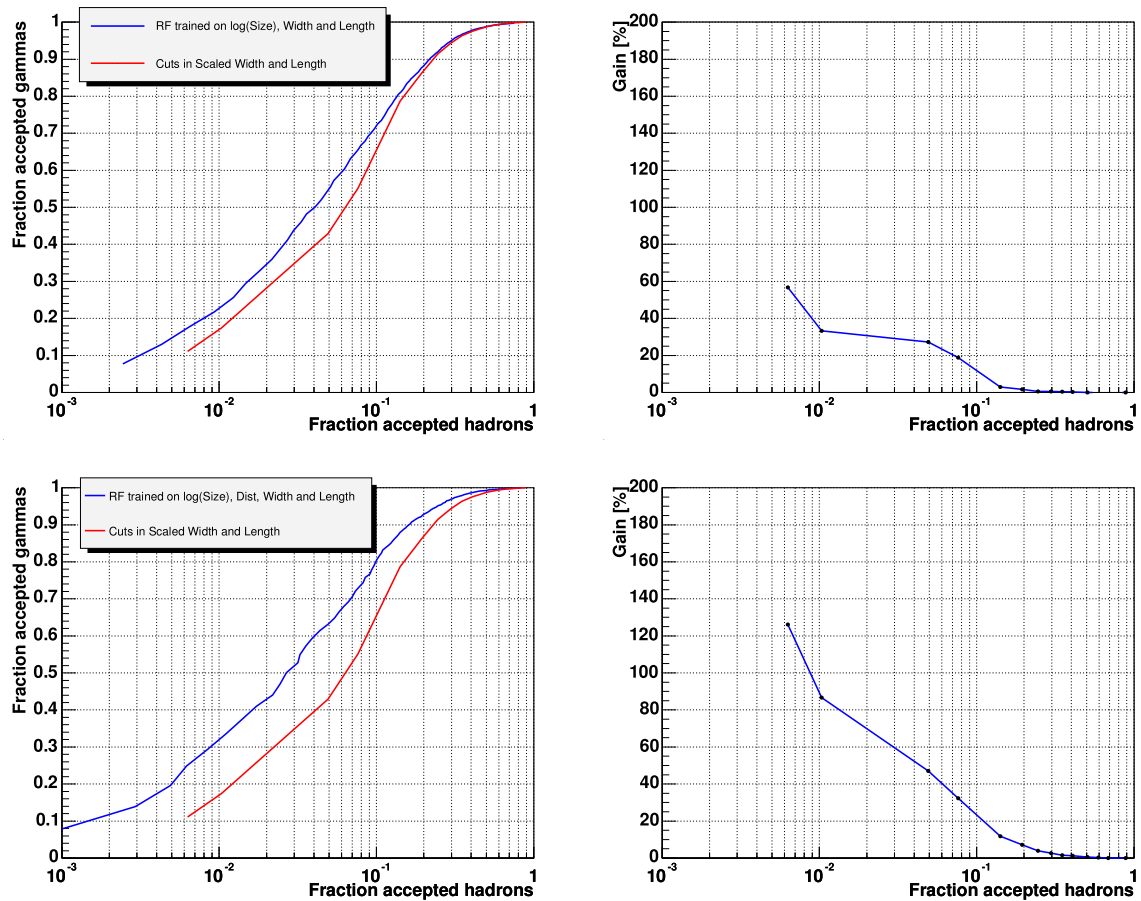
Figure 5.19: ROC curves for applying RF and Cuts in Scaled Hillas on the test sample. The parameters used in addition to the cuts in scaled Hillas are indicated by the legend. The parameters for RF are as follows: **Upper plot:** $\log_{10}(size)$, $width$, $length$. **Lower plot:** $\log_{10}(size)$, $width$, $length$, $dist$
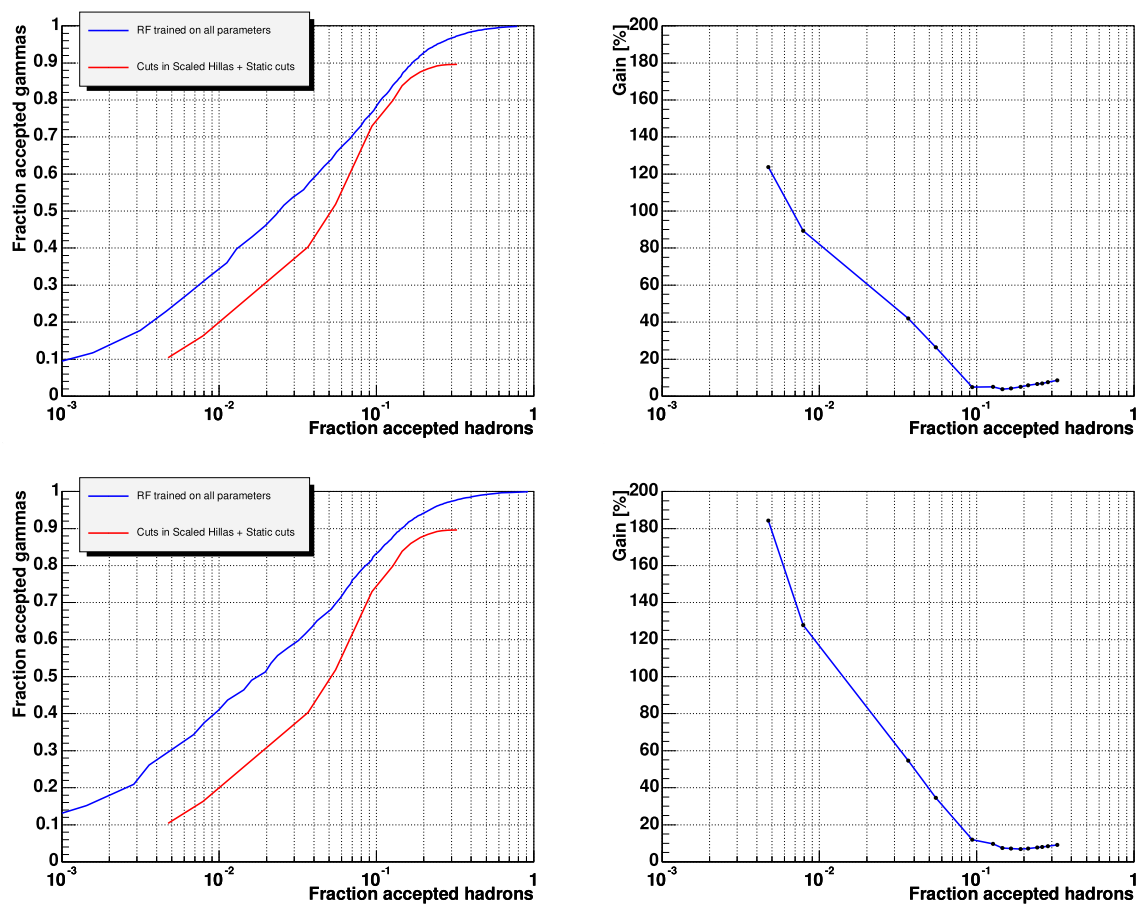
Figure 5.20: ROC curves for applying RF and Cuts in Scaled Hillas Parameters on the test sample. The parameters used in addition to the cuts in Scaled Hillas Parameters are indicated by the legend. The parameters for RF are as follows: **Upper plot:** $log_{10}(size)$, $dist$, $width$, $length$, $sign_q$, $relative\ island\ size$. **Lower plot:** $log_{10}(size)$, $dist$, $width$, $length$, $sign_q$, $relative\ island\ size$, $dens$.

For a second study also 'advanced' image parameters were taken. In a practical analysis the following cuts are used in a static manner (see also section 4.3):

- *relative island size*
  Events are accepted, if they contain less than 10 % of phe in the islands.

- $sign_q = \text{sign}(\cos\beta_q) > 0$, $\beta_q = \angle(\boldsymbol{a}_q, \boldsymbol{r}_q)$
  Only events with a charge distribution pointing towards the source position are taken. Different cut values like e.g. $\beta_q < 20°$ are not advisable since due to the correlation between β_q and *alpha* mostly the background region in the alpha plot will be cut. Therefore only the so-called 'head-tail' information contained in the charge asymmetry is exploited.

Figure 5.20 shows very significant improvements of partly more than 100 % introduced by the RF and the additional image parameters.

5.5 Application of the RF in the energy estimation

An algorithm of estimating a continuous quantity (rather than a class membership, which takes discrete values) can be constructed from the standard RF method by means of two main approaches:

- *'Forced' division into classes*
  At the beginning of this chapter 'energy estimation' was given as example of multi-dimensional event classification. One has to assign class labels to the training events according to an energy grid. As a result multiple classes are created:

$$E_0(0), E_1(1), ..., E_{n-1}(n-1) \tag{5.19}$$

  In the RF training process the related class populations are taken into account together with a more general Gini-index [BFOS83].

$$p_i = \frac{N_i}{N} \tag{5.20}$$

$$Q_{Gini} = 1 - \sum_{i=0}^{n-1} p_i^2 \tag{5.21}$$

  Here, i is the class index $0 \leq i \leq n-1$. As already shown above, the Gini index of a split is evaluated as sum of the two Gini-indices, that one obtains after the split (taking the class populations of the successor nodes).

  After the training procedure, when fully grown trees are present, the class populations inside a terminal node can be used to immediately calculate the terminal node's corresponding estimated energy:

$$E_{est} = \frac{\sum_{i=0}^{n-1} E_i \cdot N_i}{\sum_{i=0}^{n-1} N_i} \tag{5.22}$$

In the application of RF each tree returns an estimated energy and the overall mean is calculated as the final estimated energy.

- *Splitting rule based on the continuous quantity*
  It is possible to completely avoid the use of classes by introducing a splitting rule, which does not rely on class populations.

  The idea of the Gini-index (with its interpretation as binomial variance of the classes) as split rule is a purification of the class populations, i.e. a separation of the classes, in the subsamples after the split process. Similarly, when using the variance in energy as split criterion, the subsamples are purified with respect to their energy distribution.

$$\sigma^2(E) = \frac{1}{N-1} \sum_{i=1}^{N} (E_i - \overline{E})^2 = \frac{1}{N-1} \cdot \left[\left(\sum_{i=1}^{N} E_i^2 \right) - N \cdot \overline{E}^2 \right] \tag{5.23}$$

In analogy to the Gini-index of the split, the 'variance' of the split is calculated by adding the 'subsample energy variances' taking into account the node populations as weights:

$$\sigma^2(E) = \frac{1}{N_L + N_R} (N_L \sigma_L^2(E) + N_R \sigma_R^2(E)) \tag{5.24}$$

5.5.1   Performance of the RF energy estimation

In the following the results of the RF energy estimation are presented using the Crab Nebula-like gamma Monte Carlo sample, which was already described in section 5.4. But now the energy range of $10\,\mathrm{GeV} < E < 30\,\mathrm{TeV}$ is taken since there is no need for adopting to a proton MC. The following quality cuts were imposed on this sample:

- Static dist cut: $dist > 0.3^\circ$

- Leakage cut: $leakage < 0.1$

These cuts remove events, which provide only a weak basis for an energy estimation, since the *size-energy* and *dist-impact parameter* dependences become wide-spread if exceeding the cut limits (see below for further explanations).

Let E_{true} and R_{true} denote the true (Monte Carlo) energy and the true impact parameter respectively. Figures 5.21 and 5.22 show the dependences $log_{10}(size)$-$log_{10}(E_{true})$ and $dist$-R_{true}. The strong energy-*size* dependence is the basis for any energy estimation. Yet, since the distribution of the Cherenkov photons inside the Cherenkov light pool is not completely constant and changing with the distance between telescope and shower axes (the impact parameter), an estimation of the impact parameter provides important

information for a correction of the first-order energy-*size* dependence. This is the reason, why in the case of Cherenkov telescope arrays usually a superior energy resolution is achieved: for a telescope operation in coincidence mode the impact parameter can be very well estimated from the cross section region of the event's major image axes as recorded by the telescopes of the array.
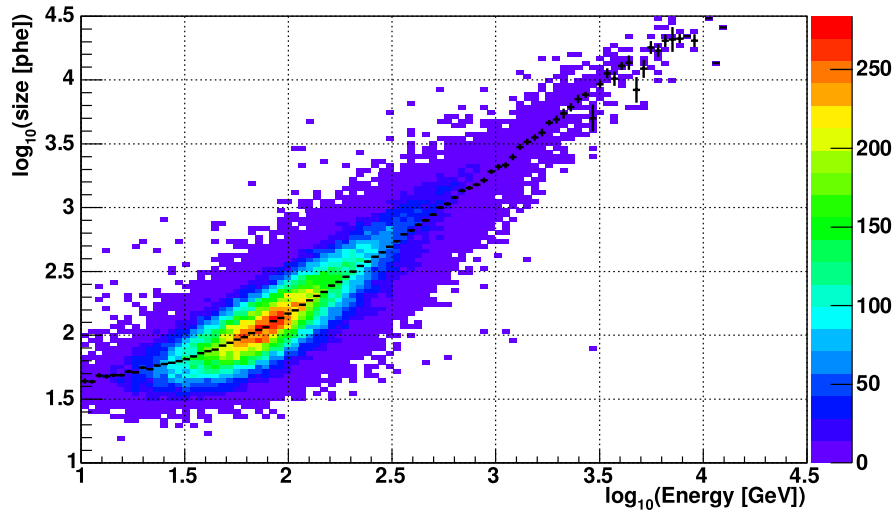
Figure 5.21: $\log_{10}(size[\text{phe}])$ versus $\log_{10}(E_{true}[\text{GeV}])$. A linear dependence (as first-order approximation) is visible reflecting the approximative proportionality between energy and photon density inside the Cherenkov light-pool.

Thus, it is advisable to additionally use image parameters, which are correlated with the impact parameter (first of all the *dist* parameter), as input for the RF energy estimation.

The training and test of the Random Forest (RF) energy estimation method was done splitting the gamma sample into 80 thousand events for training and 80 thousand for the testing. As for the RF settings, already after 30 trees no significant change in the RF performance (in particular the energy resolution) could be noticed. Node size and number of trials were set to the 'best values' (see section 5.2).

Classification and regression versions of RF (as described at the beginning of this section) showed approximately agreeing results, so that in the following merely the performance of the RF regression method will be shown.

In case of the classification version, the energy grid reaching from 10 GeV up to 30 TeV (the energy range covered by the MC simulation) was split into 200 logarithmically equidistant bins (later on it was noticed that already 100 would have been sufficient). In a first approach the following 'standard' image parameters were taken as input for RF (for a definition of the image parameters see above and section 4.3):

- *size, dist, width, length*

- *dens*
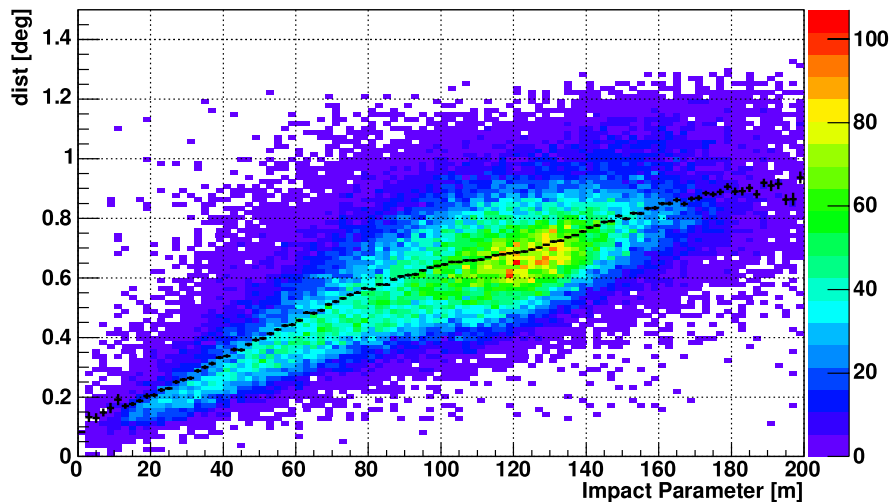
Figure 5.22: Hillas Parameter *dist* [deg] versus impact parameter R_{true} [m]. An approximately linear correlation is apparent, which can be understood from the shower and telescope geometries.

- *leakage*

- $\cos\theta$ with θ being the zenith angle

The results of the RF energy estimation can be seen from figure 5.23. The estimated energy shows a good proportionality with the true energy down to at least 50 GeV ($\log_{10}(E/\text{GeV}) = 1.7$) keeping the energy resolution at still reasonably good values $\lesssim 30\%$. The bias is mostly below the 10 % range. For energies exceeding 1 TeV the bias is as low as 7 % and the energy resolution of 22.2 % is well comparable with previous generation IACTs: In the case of the HEGRA CT array best values of 23 % for the energy resolution were reached [Sch01].

In a further approach also the charge and time asymmetries were taken into account (the corresponding vectors are $\boldsymbol{a_q}$ and $\boldsymbol{a_t}$, see section 4.3), since an improvement of the energy estimation is expected according to figures 5.24 and 5.25: One can clearly see a correlation of $|\boldsymbol{a_q}|$ and $sign(\cos\beta_t)|\boldsymbol{a_t}|$ with energy and impact parameter respectively. In the case of the time asymmetry the sign of the asymmetry angle must be taken into account, since there is a swapping in the temporal development of the Cherenkov light image in the camera (the so-called seagull effect, see section 4.3.1)

The additional usage of the charge asymmetry vector length $length_q = |\boldsymbol{a_q}|$ does not improve the energy resolution (see figure 5.26). This means that the correlation as shown in figure 5.24 is too weak and/or already exploited by the normal *size* parameter. The result of the RF energy estimation including the charge *and* time asymmetries can be seen in figure 5.27. The apparent difference in comparison with figure 5.23 is that the energy resolution significantly improved from about 22.2 % to 18.5 % for energies $E > 1$ TeV.
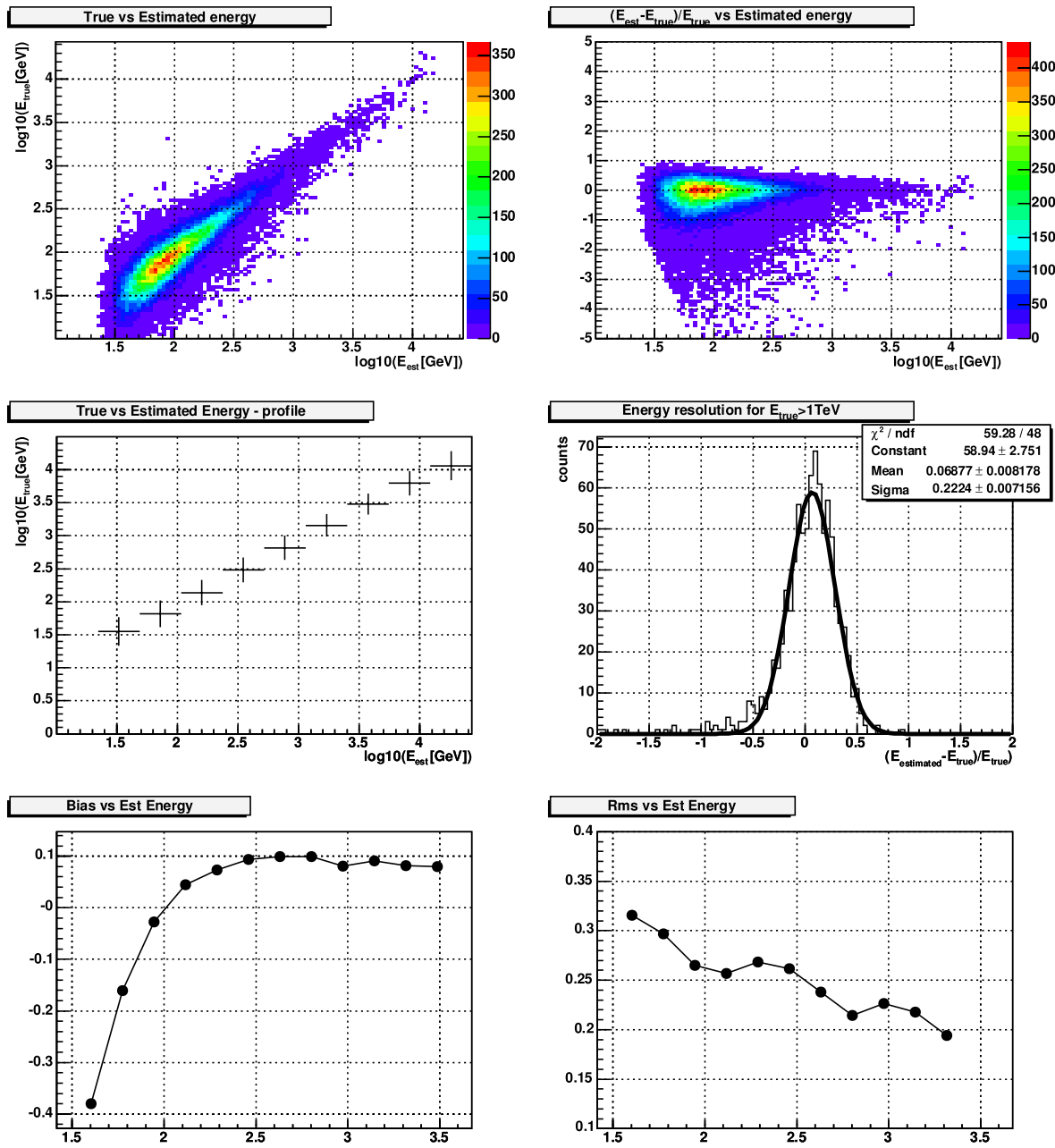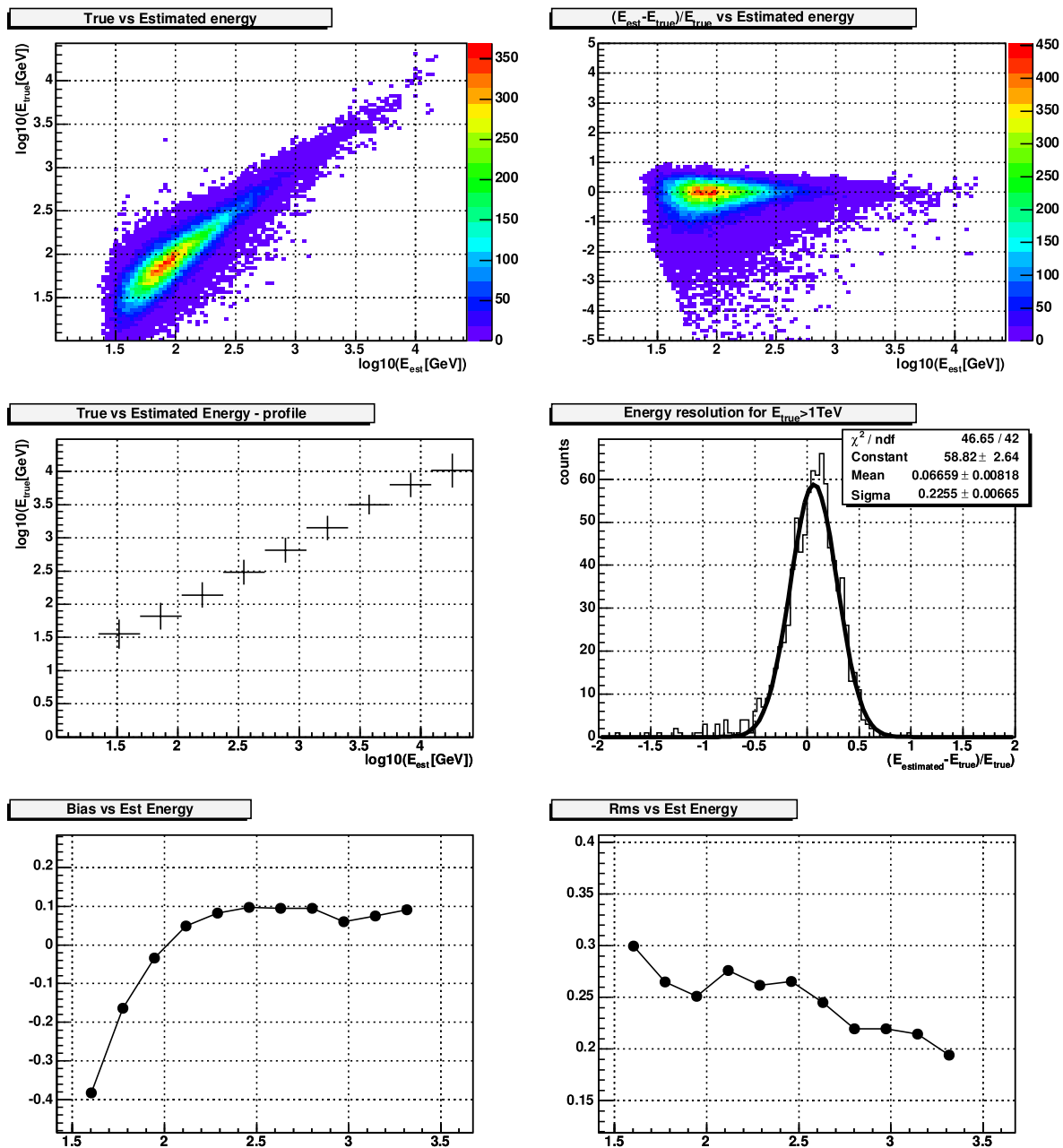
Figure 5.23: Result of the RF energy estimation applied to the test sample. Only the 'standard image parameters' described in the text are used. **Top left:** True energy E_{true} versus estimated energy E_{est}. **Top right:** Energy resolution $\frac{E_{est}-E_{true}}{E_{true}}$ as function of estimated energy E_{est}. The asymmetry of the energy resolution in particular at lower energies is due to a similar asymmetry of the *size*-distribution (small *size* values appear more frequently). **Middle left:** Bin wise mean of true energy E_{true} versus estimated energy E_{est}. **Middle right:** Energy resolution obtained for $E_{true} > 1\,\mathrm{TeV}$. **Bottom left:** Bias versus estimated energy. **Bottom right:** RMS of energy resolution (Gaussian fit) versus estimated energy.
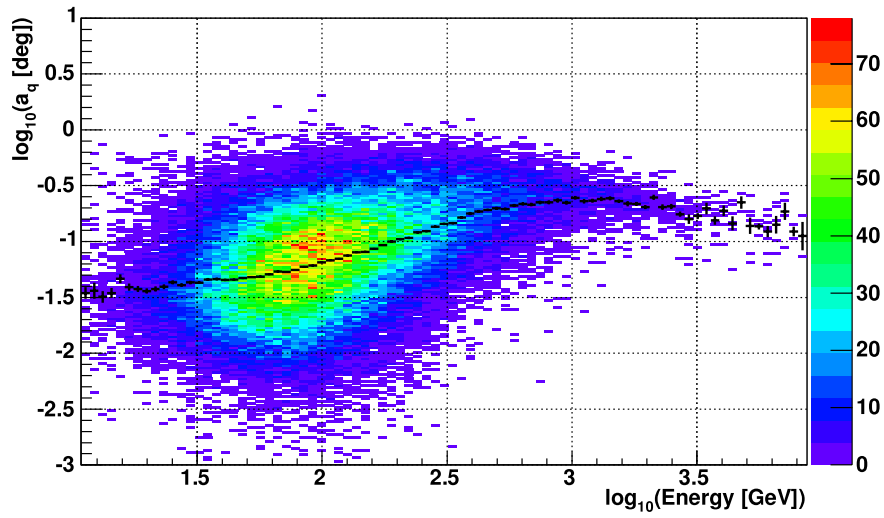
Figure 5.24: Dependence of $|\boldsymbol{a_q}|$ [°] (length of charge asymmetry vector) on true $\log_{10}$-Energy [log_{10}(GeV)].
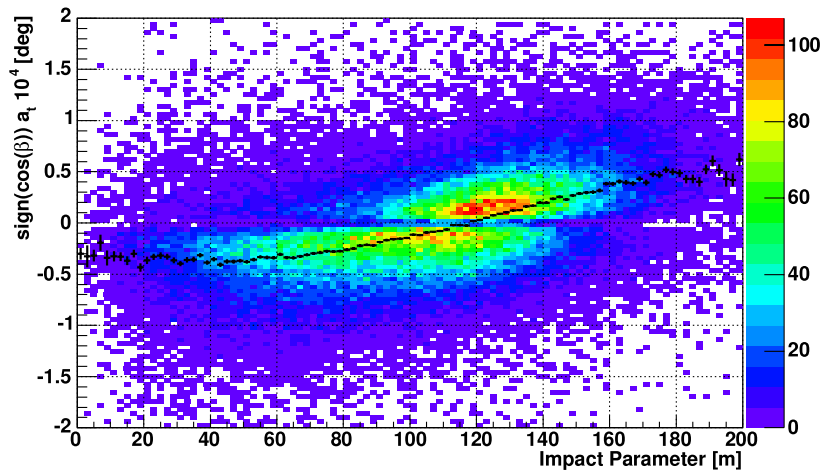
Figure 5.25: Dependence of $sign(\cos\beta_t)|\boldsymbol{a_t}|$ [°] on true impact parameter R [m]. The region around $|\boldsymbol{a_t}| = 0$ is sparsely populated since arrival time fluctuations ($\rightarrow$ time resolution) in the Cherenkov light image make a time asymmetry vector length of zero very unlikely.

Figure 5.26: Result of the RF energy estimation applied to the test sample. Now, additionally the charge asymmetry parameter $length_q$ was taken into account. For an explanation and comparison of the plots, see figure 5.23. The energy resolution is not significantly changed.
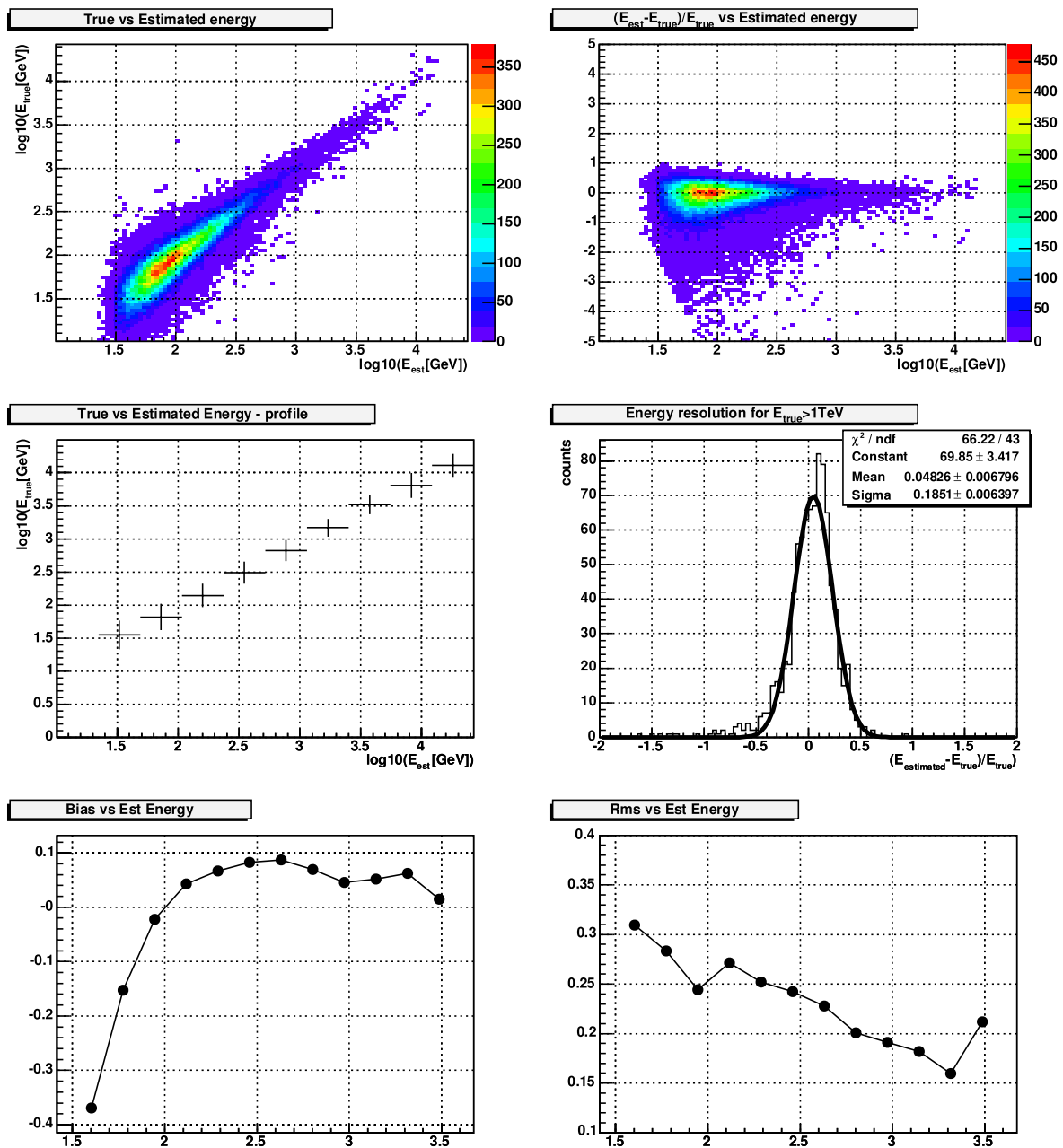
Figure 5.27: Result of the RF energy estimation applied to the test sample. Now, additionally the charge and time asymmetries were taken into account. For an explanation and comparison of the plots, see figure 5.23. The energy resolution is significantly improved from about 22.2 % to 18.5 %.

The improvement in energy resolution is also notable for low energies $E \approx 100\,\mathrm{GeV}$ where now values around 25 % are achieved.

It shall be noted that the values mainly of the energy resolution depend on the precuts or quality cuts. The quality cuts as introduced at the beginning of this section are the same as those ones later used in the analysis of Crab Nebula and PG1553+113 observational data.
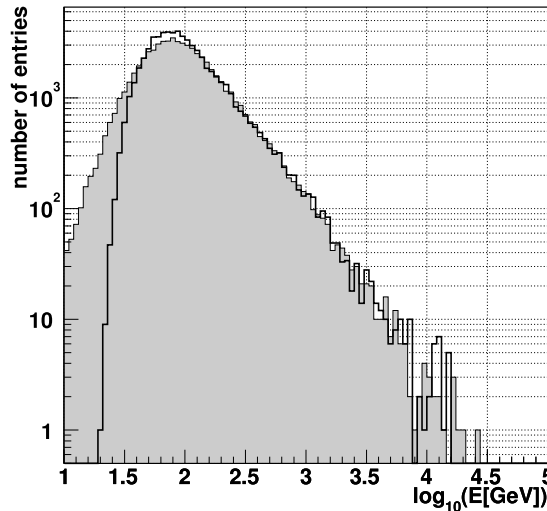
Figure 5.28: Comparison of true simulated MC spectrum (grey shaded histogram) and reconstructed spectrum (black line). Significant distortions of the reconstructed energy spectrum are limited to estimated energies $E \lesssim 100\,\mathrm{GeV}$.

A further important consideration is the reconstruction quality concerning the energy spectrum. As can be seen from figure 5.28 significant distortions of the reconstructed energy spectrum are restricted to energies $E \lesssim 100\,GeV$. Figure 5.29 shows reconstructed energy spectra with a binning similar to that one, which will be used later on in the source analysis. Two different Monte Carlo simulations are tested here: the first one shows a Crab Nebula-like energy spectrum with power-law (or photon-) index $\Gamma = 2.6$ and the second one shows a power-law index of $\Gamma = 4.0$ (based on the measured differential energy spectrum of PG1553+113 in [A+06b]). As can be seen from the histograms in figure 5.29 both spectra are well reconstructed in the energy interval $100\,\mathrm{GeV} < E < 1\,\mathrm{TeV}$, which covers the range of significant excess in the case of PG1553+113 observations (see chapter 7).

Photon index Γ = 2.6

Photon index Γ = 4.0
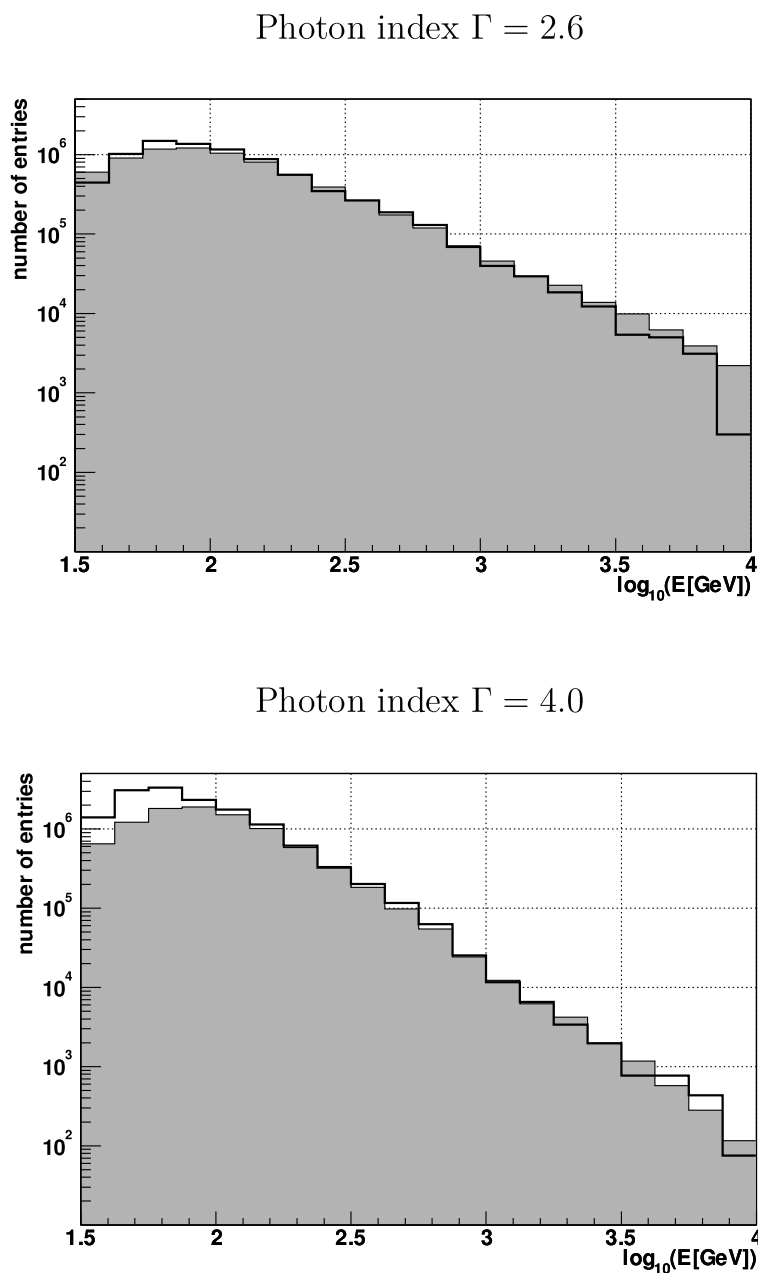
Figure 5.29: Comparison of true simulated MC spectrum (grey shaded histogram) and reconstructed spectrum (black line) for two simulations, which show energy spectra with photon indices as indicated. No distortions can be seen in the energy interval $100\,\text{GeV} < E < 1\,\text{TeV}$, which is decisive in the later analysis of PG1553+113.

5.6 Implications on the Random Forest Method

The application of the Random Forest method both in the g/h-separation and the energy estimation lead to the following conclusions:

- Though RF is not an analytical method, it produces *stable results*, i.e. a repeated training process (different seed values for the tree randomization) yields identical performances. In the case of energy estimation even the splitting rule (Gini index in the classification and energy variance in the regression variant of the RF energy estimation) does not seem to induce significant changes.

- *Automization of g/h-separation and energy estimation*
  There are only a few tunable parameters, which can be easily chosen according to simple criteria. The task of the user is to provide proper training samples (obviously, special care must be taken for a good agreement between Monte Carlo simulations and real data).

- *Fast training and classification*
  Benchmarks from 1.5 GHz CPU (Athlon XP), training and test sample containing each 10000 events, 100 trees used for classification, each tree completely grown (node size 1), total of 10 image parameters used, 3 trials in random split selection

  Training process: 1 min.

  Application process: 20 sec.

  A comparable analysis technique like Neural Networks generally demand a more time consuming training process.

- Random Forest can take *any numerical input parameters*, which can have strong correlations. In particular, it is very easy to test and utilize new parameters (e.g. charge and time asymmetries for RF energy estimation).

A drawback, which holds for any advanced g/h-separation method requiring a training process, is that one relies on a good Monte Carlo Simulation for the gammas.

Since electromagnetic interactions are well known (in contrast to a more complicated hadronic shower simulation), the simulation of gamma induced air showers is highly reliable. The remaining problem is a careful treatment of the atmosphere and the detector itself. At this point one must make simplifications and test if these perform to expectation.

Chapter 6

Measurement of the VHE γ-ray spectrum from the Crab Nebula

6.1 Overview

The Crab Nebula is the most frequently studied object in the gamma-ray sky. It has been observed in the VHE energy range from about 100 GeV to 50 TeV. The unpulsed emission has been measured by the EGRET experiment in the HE energy range up to about 10 GeV (see also section 1.1).

The HE gamma-rays are believed to be produced by inverse Compton (IC) scattering of relativistic electrons on low-energy seed photons [O⁺01]. These seed photons may be synchrotron photons or external photons originating from a combination of the cosmic microwave background and a local contribution from dust and starlight (see section 1.5.1 and figure 1.9). In the case of the Crab Nebula, the HE emission is well explained as IC scattering of relativistic particles on those photons, which have been emitted at lower energies as synchrotron radiation by the same particles [GS06]. For other pulsar wind nebulae much lower synchrotron luminosities are observed as compared to the Crab Nebula, which implies that the seed photons for IC scattering are in these cases primarily of external origin [GS06].

Thus, for the Crab Nebula the following IC model scenario arises: Electrons are accelerated at the pulsar wind termination shock and, propagating through the nebula, they produce both synchrotron and inverse Compton emission. Also pulsed gamma-emission, assumed to be produced by the pulsar, has been seen by EGRET, but not (yet) by ground based experiments in the higher energy regime. Polar cap models predict sharp cutoffs near 10 GeV whereas outer gap models predict pulsed emission at energies as high as 50 GeV [O⁺01]. Therefore a detection of pulsed emission at energies $E \gtrsim 50$ GeV would rule out current polar cap models (see also the item *pulsars* in the list of main targets for MAGIC observations, chapter 3).

The Crab Nebula is often referred to as the 'standard candle' of gamma-ray astronomy due to its (unpulsed) steady IC emission. This steady VHE γ-emission is, as mentioned above, well known in the energy range from about 100 GeV up to several TeV. Thus, the Crab Nebula energy spectrum can be used to check the reliability of new analysis techniques. The spectrum shown in figure 6.1 was derived from observations of the Crab

Nebula in 2004 (red marked data points) and 2005 (blue marked data points) [WM05]. The reader is also referred to [A$^+$00, MBFG05] for further information about the well-established VHE γ-emission of the Crab Nebula. A subset of ~ 1.7 hours of the 2004 data is taken as 'touchstone' for the analysis presented in this work.
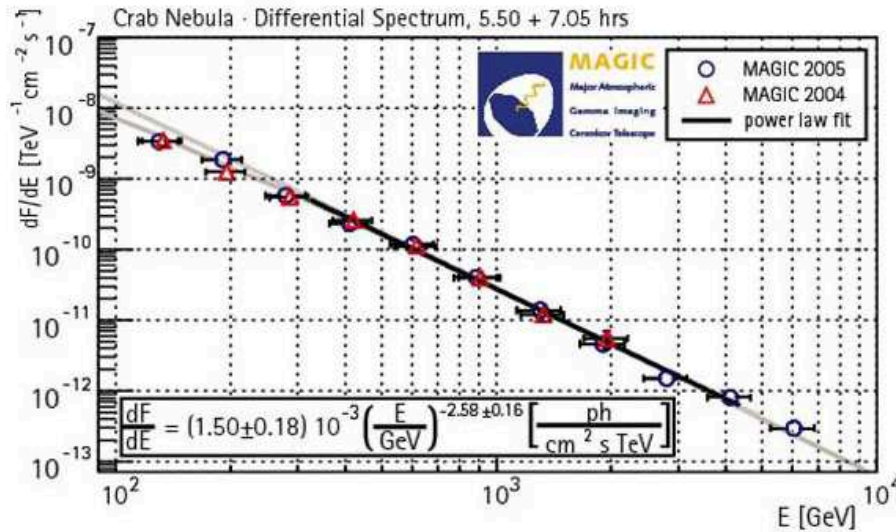

Figure 6.1: The VHE γ-ray energy spectrum of the Crab Nebula as measured with the the MAGIC Telescope in 2004 and 2005 [WM05]. The faint grey lines visible in the energy range below 300 GeV are extrapolations of the Crab spectra measured by the HEGRA [A$^+$04] (upper line) and Whipple [H$^+$98] (lower line) experiments.

A spectrum covering the energy range from below 100 GeV up to 1 TeV will be determined and checked for consistency with the just mentioned independent analysis. Furthermore, the influence of different g/h-separation settings on the measured spectrum will be studied.

These studies on the γ-ray spectrum of the Crab Nebula are intended to prepare the analysis of the VHE γ-ray emission of PG1553+113 presented in the next chapter. As will be shown later, significant γ-ray emission was detected for the BL Lac source PG1553+113 in the energy range $100\,\mathrm{GeV} \lesssim E \lesssim 500\,\mathrm{GeV}$.

6.2   Consistency between Monte Carlo simulations and observational data

After analyzing the observational data according to the methods described in chapter 4 an event sample containing image parameters is at disposal for further analysis. At this stage the adjustment of parameters in the MC simulation can be performed. For this purpose, different settings in PSF and reflectivity of a muon MC (μ-MC) simulation are tested. Figures 6.2 and 6.3 show the distributions of *size* vs *arc radius* and *arc width* respectively. These plots were produced from a μ-MC and observational data selecting
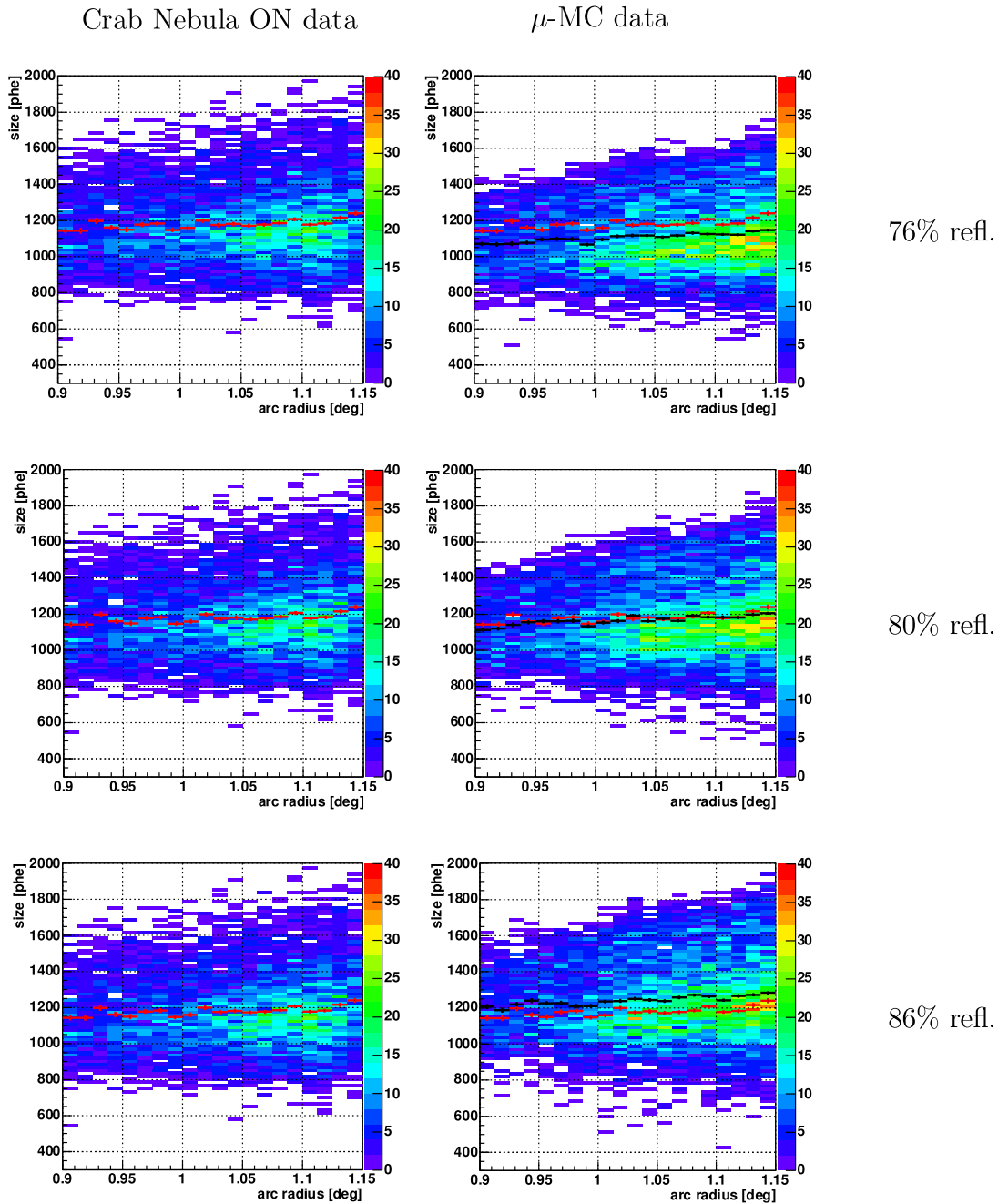
Figure 6.2: **Left:** *arc radius* vs *size* diagram for μ-rings selected from Crab Nebula ON data (see section 4.3.2 for more explanations). The *size*-dependent mean value is plotted as red profile. **Right:** *arc radius* vs *size* diagram for μ-rings selected from different μ-MC sets with reduced reflectivities 76 %, 80 % and 86 % (from top to bottom). The total reflectivity is obtained by multiplying the reduced reflectivity value with the wavelength dependent reference reflectivity (86 %-90 %), see also [BM04]. The black curve shows the profile plot. For comparison the (red) profile plot from Crab Nebula ON data is overlaid.

$+\,7\,mm$ $+\,10\,mm$
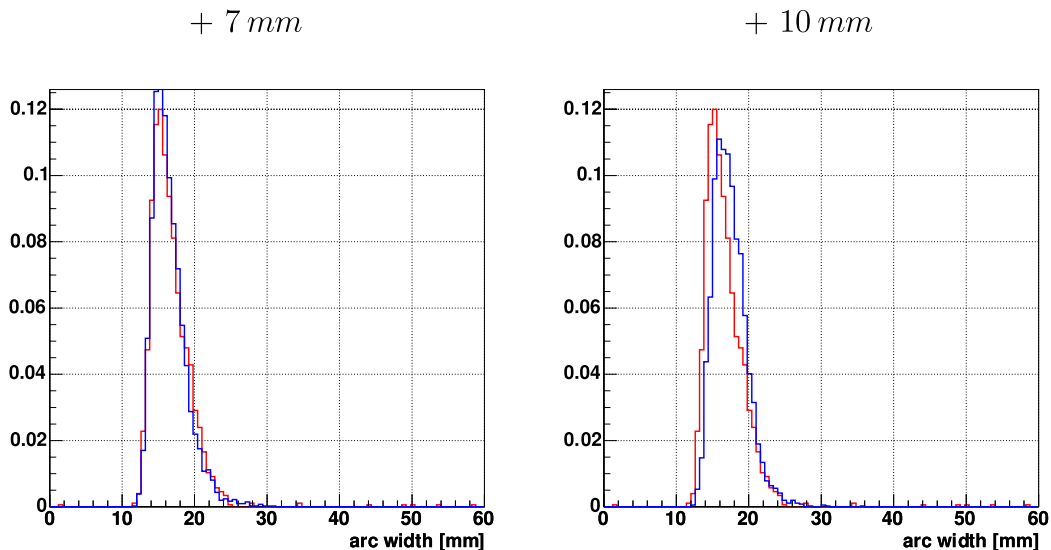
Figure 6.3: Blue histograms: Muon *arc width* distributions for the additional PSF values 7 mm (left) and 10 mm (right). Red histogram: *arc width* distribution as determined from muons selected from Crab Nebula ON data. The total PSF is obtained by adding in quadrature the default value of 7 mm yielding about 10 mm for the 'best matching' MC setting, which corresponds to 0.034°.

'good' muons by demanding a high χ^2-probability of the ring fit and requiring at least half-rings (see section 4.3.2). By comparing diagrams obtained from Crab Nebula ON data and from different μ-MC sets, one can choose the 'matching' reflectivity (the plot shown in figure 6.2 is not changed significantly for different reasonable PSF settings in the range $6-15\,$mm) and thereafter the 'matching' PSF.

In order to quantify the 'degree of mismatch', χ^2 values for the comparison of the histograms under consideration (see figures 6.2 and 6.3) were evaluated. The μ-MC yielding the smallest 'degree of mismatch' was taken as 'best matching'. Using this approach the reflectivity and PSF values of 71 % (maximum of the wavelength-dependent reflectivity) and 10 mm, respectively, are obtained in the case of Crab Nebula ON data. The PSF of 10 mm corresponds to 0.034°, which is in good agreement with [GM05], where a PSF of $0.037 \pm 0.006^{\circ}$ is deduced from observations in September 2004.

Figures 6.5 to 6.7 and A.1 to A.7 (see appendix A) show a comparison of the distributions of all image parameters relevant for the subsequent analysis, for Crab Nebula ON data and an adjusted proton MC. An alpha-cut $|alpha| > 20^{\circ}$ was imposed in order to reduce the γ-ray content of the ON data. The OFF data are not strictly adequate to check the MC agreement, since they are usually taken non-contemporaneously. OFF data are taken on sky regions, selected in such a way that NSB and zenith angle range are similar to the ON observations. Nevertheless, there may be systematic differences in the background conditions of ON and OFF data. Furthermore, the ON data enter into the training of the g/h-separation (the RF is able to deal with 'impure' training samples). Therefore, it is preferable to check the MC consistency using ON data.
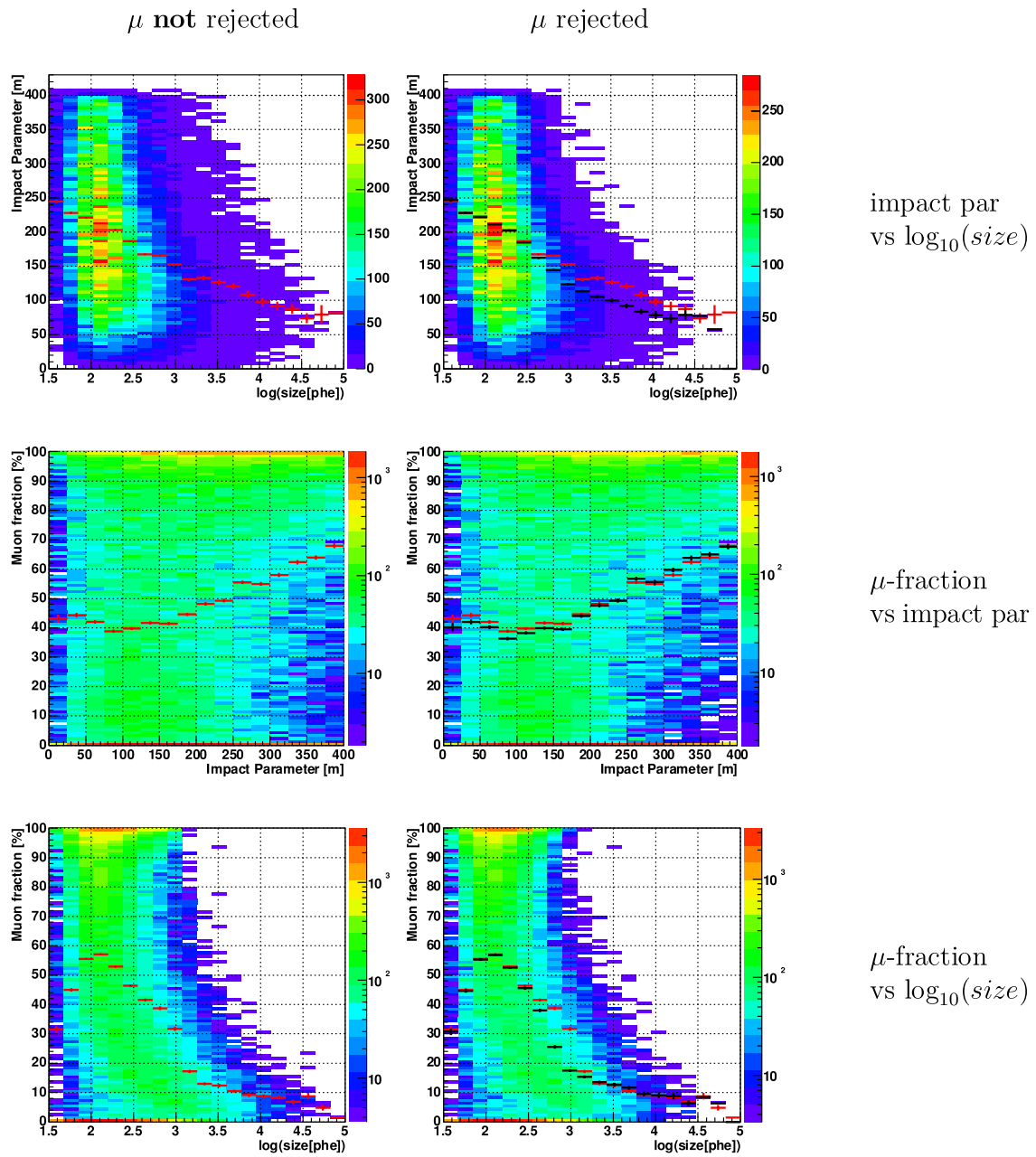
Figure 6.4: Plots derived from a proton-MC without (left side) and with μ-rejection (right side) as explained in the text. The so-called 'muon fraction' denotes the percentage of Cherenkov light hitting the camera, that was produced by muons. **Top:** Impact parameter R vs $\log_{10}(size)$. **Middle:** Muon fraction as function of impact parameter. **Bottom:** Muon fraction as function of $\log_{10}(size)$.
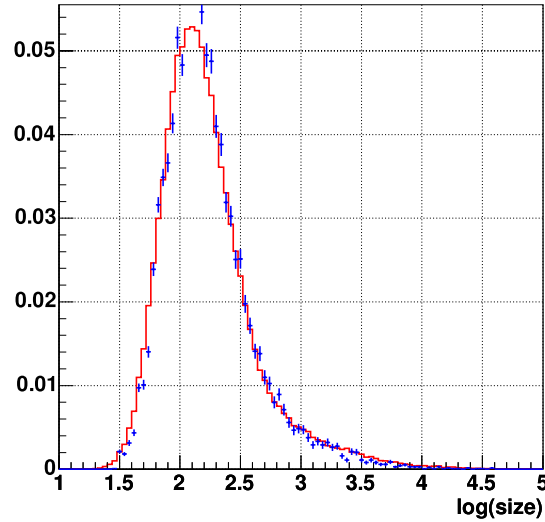
Figure 6.5: Comparison of the *size* distributions of MC protons (blue data points) and Crab Nebula ON data (red histogram) showing good agreement. A strong muon filter has been applied (see text).

An 'inconvenience' arises from the fact that the standard proton MC (p-MC) simulation considers impact parameters R only up to $R = 400$ m. However, at $R = 400$ m the fraction of triggering protons is *not* negligible and *not* approaching zero (see figure 6.4). The p-MC simulations show, that the muon fraction increases strongly with increasing impact parameter (figure 6.4, middle): The events triggering at high R are mostly muons. Thus, for both ON data and p-MC samples high χ^2 values of the muon ring fit as well as a rejection of half-rings were demanded in order to reduce and possibly harmonize the muon contributions in the samples. However, due to the large impact parameter of the parent shower these muons, having most probably also big impact values R_μ, show up in the camera as arc-fragments with rather small *size* values. The above mentioned muon ring rejection provides a satisfying muon reduction in the *size* region $\log_{10} size \gtrsim 2.5$ (see figure 6.4), which can therefore be taken for comparison with observational data. Furthermore, for all following plots of the type $p(\log_{10}(size))$ (with p being the image parameter under consideration plotted as function of $\log_{10}(size)$) a Monte Carlo 'test plot' is shown with the $p(\log_{10}(size))$ distribution for $R < 200$ m. If the (unknown) contribution of triggering events with $R > 400$ m introduces significant changes, a comparison of the plots related to $R < 200$ m and $R < 400$ m should already indicate the region in the $p(\log_{10}(size))$-histograms, which is most strongly affected by 'big R events' and therefore not reliable to judge a MC agreement.
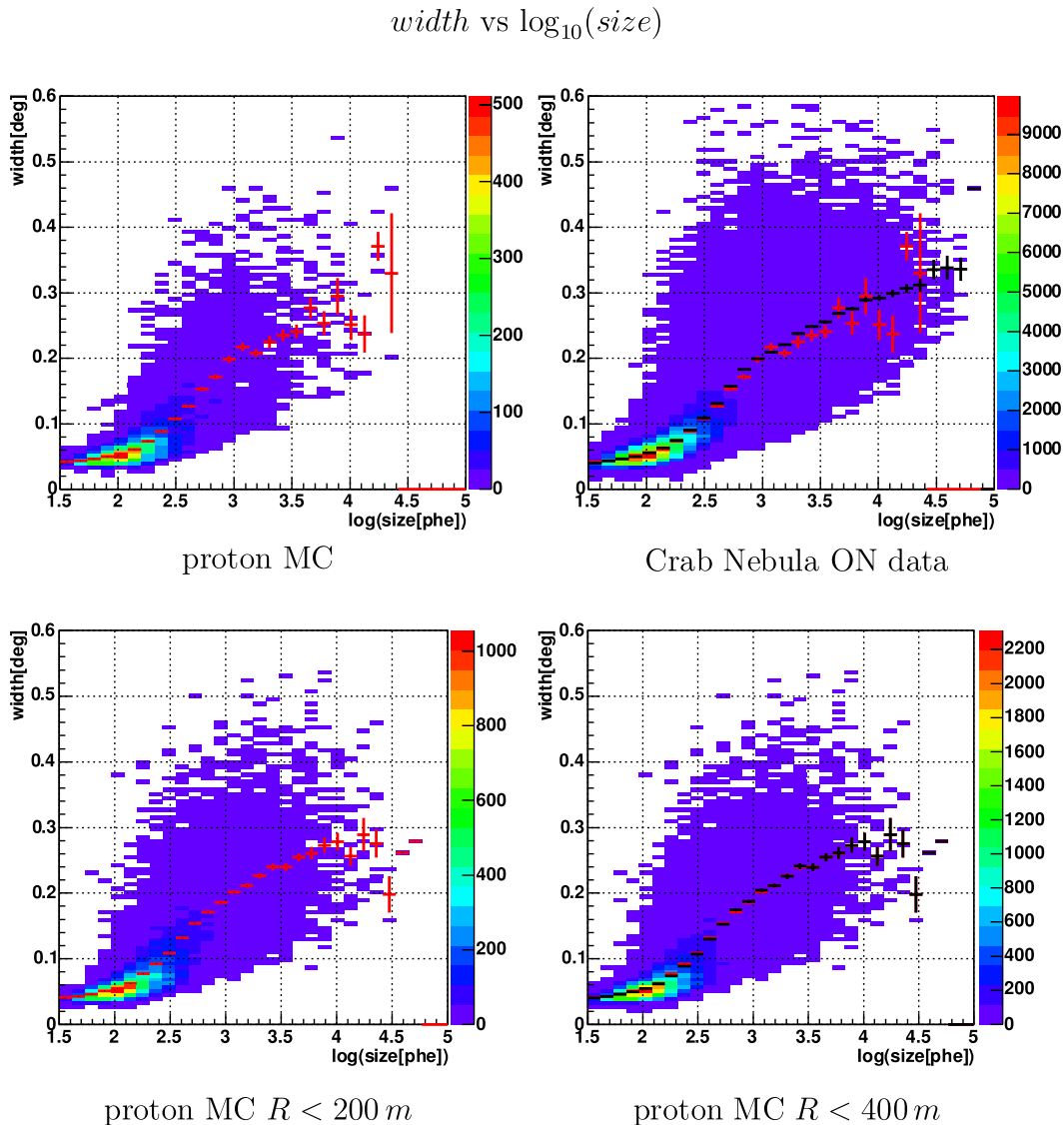
$width$ vs $\log_{10}(size)$

Figure 6.6: $width$ versus $\log_{10}(size)$ histograms. **Top:** $width$ vs $\log_{10}(size)$ plot for the p-MC (histogram on the left and red profile plot) and Crab Nebula ON data (histogram on the right and black profile plot). **Bottom:** $width$ vs $\log_{10}(size)$ plot of the p-MC for impact parameters $R < 200$ m (histogram on the left and red profile plot) and for impact parameters $R < 400$ m (histogram on the right and black profile plot). The region $2 < \log_{10}(size) < 3$ is slightly sensitive to the impact parameter range of the MC simulation. A comparison of the top left and top right histograms and profile plots can therefore be done only with reservations. However, different MC settings introduce a significant worsening of the MC agreement (see next figure).

$width$ vs $\log_{10}(size)$

proton MC add. PSF $= 10\,mm$ Crab Nebula ON data

proton MC add. PSF $= 20\,mm$ Crab Nebula ON data
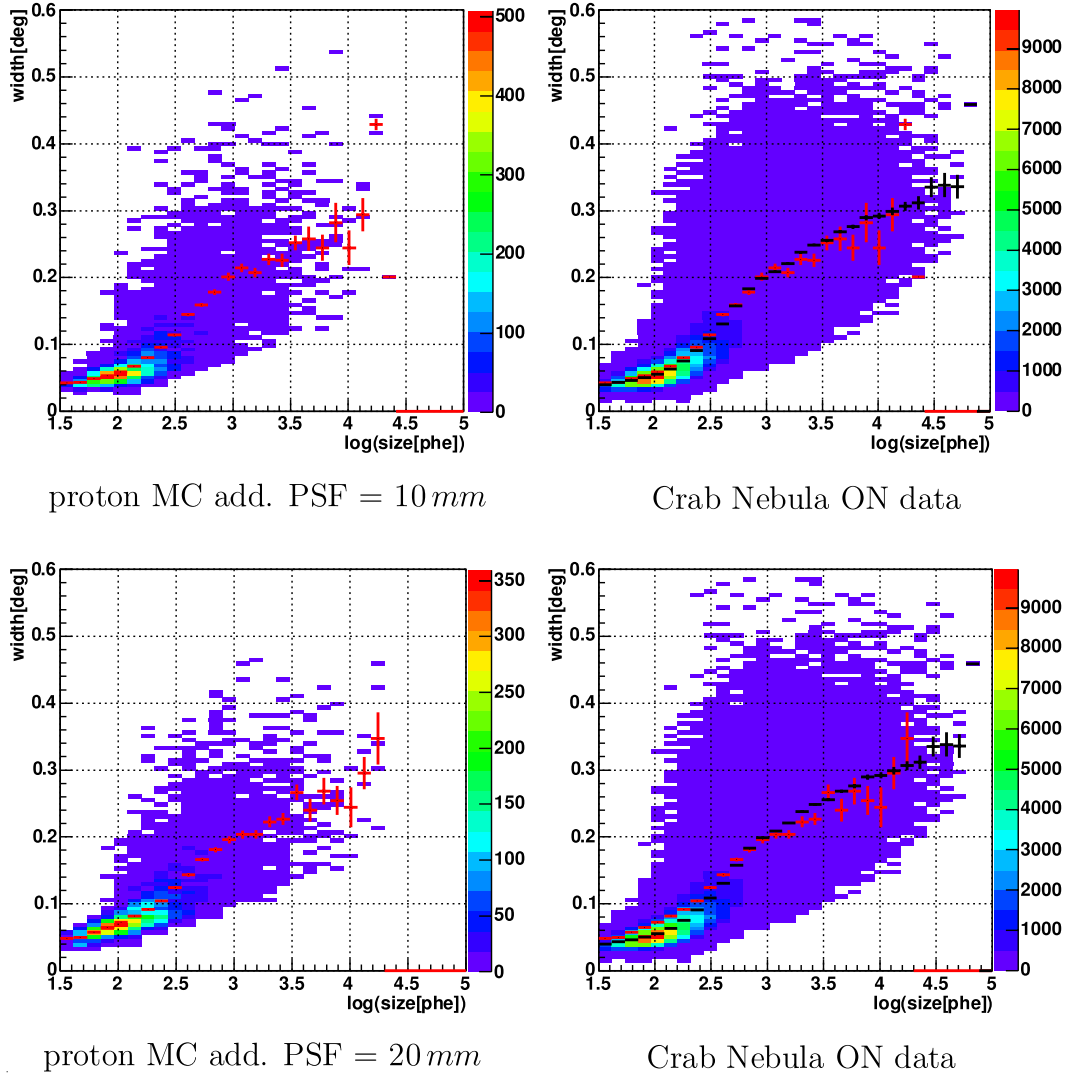
Figure 6.7: $width$ versus $\log_{10}(size)$ histograms for different MC settings (compare with figure 6.6, where a PSF of 7 mm is used). A Monte Carlo mismatch for the $width$ distribution at small values of $\log_{10}(size) \lesssim 2.5$ becomes visible, especially for the MC with PSF=20 mm. The reason is that the lateral spread in particular of small images resembles the PSF.

6.3 Optimization of g/h-separation

As a preparation for the g/h-separation using the RF, training samples of gammas and hadrons must be given. According to the discussion in chapter 5 a favorable choice of these training samples is as follows:

- gamma-sample: Monte Carlo gammas

- hadron-sample: Crab Nebula ON data

As already mentioned above, the OFF data never exactly match the observation conditions under which the ON data were taken. It is therefore advantageous to directly use data from ON source observations. The performance of the RF is not significantly reduced even if the gamma-ray content of the ON data is as high as 1 %-5 % (see chapter 5, section 5.4). For a higher gamma-ray fraction a deterioration of the performance is expected, yet γ-ray fluxes are determined correctly: The worsening in performance introduced by contaminated training data (impure training data) affects the gammas in the MC gamma test sample and the gammas contained in the analyzed ON data in the same way, which results in correct cut efficiencies and effective collection areas.

The hadron training sample is composed of events randomly extracted from the ON data using weights in order to obtain a $\cos\theta - \log_{10}(size)$ distribution matching that of the gamma MC (γ-MC) sample. Both samples of gammas and hadrons contain 20000 events, which is sufficient to provide a good basis for the g/h-separation (see section 5.4). Two sets of image parameters are used for the g/h-separation:

- Set 1: 'Standard' image parameters:

    - $\log_{10}(size)$, *dist*
    - *width*, *length*
    - $\cos\theta$

- Set 2: All parameters from set 1 and the following additional image parameters:

    - *dens*
    - *sign$_q$*
    - *relative island size*

Set 2 is expected to yield a better performance of the RF (see section 5.4). For the definition of the parameters see section 4.3.1. Figures 6.8 to 6.10 show the *hadronness vs size*, *Q-value vs size* and *alpha vs hadronness* diagrams.

In view of the weak γ-emission from the BL Lac PG1553+113 presented in the next chapter the g/h-separation is optimized for weak sources. Equation 4.48, which actually underestimates the significance [GH05], can easily be evaluated for the two cases of a very weak and a very strong source. One obtains the following asymptotic behaviour of the significance S
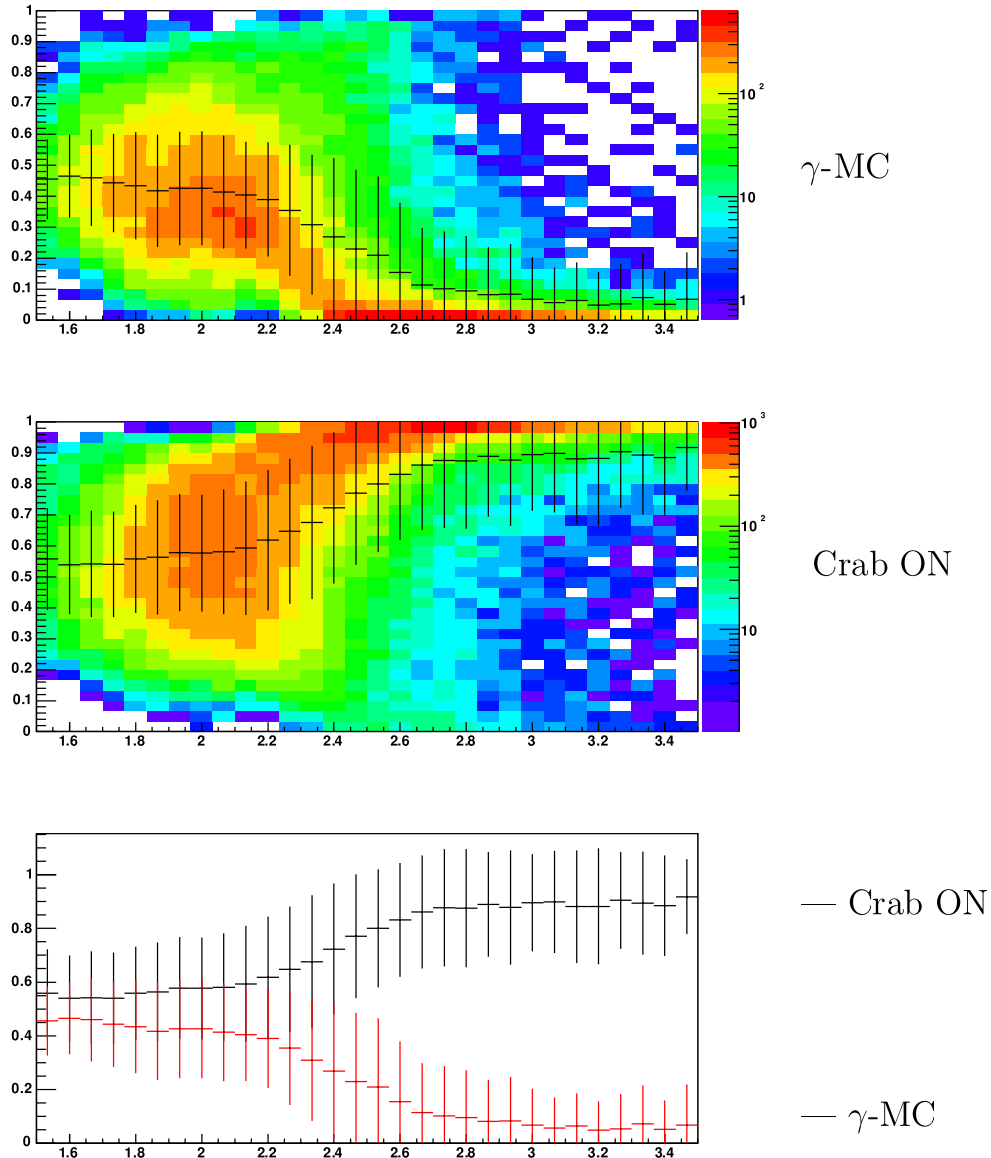
Figure 6.8: Hadronness h versus $\log_{10}(size)$ for a RF training using the image parameter set 2 (see text). The plots have been created from test samples (no overlap with the training sample). **Top:** MC gammas **Middle:** Crab ON data. Probably, a small γ-contribution is visible for $2.4 < \log_{10}(size) < 3$ and $h < 0.1$. **Bottom:** The two profile plots from MC gammas and Crab ON data overlaid in one plot. The vertical error bars denote the variance of the event distribution within a *size* bin and thus do not show the error of the mean (i.e. the error of the data point). Below $\log_{10}(size) \approx 2.7$ (500 phe) the average hadronness of gammas and hadrons start to approach each other, until below $\log_{10}(size) \approx 2.3$ (200 phe) a good separation is hardly possible.

Q-value (from γ-MC)   Efficiency ε_γ (from γ-MC)
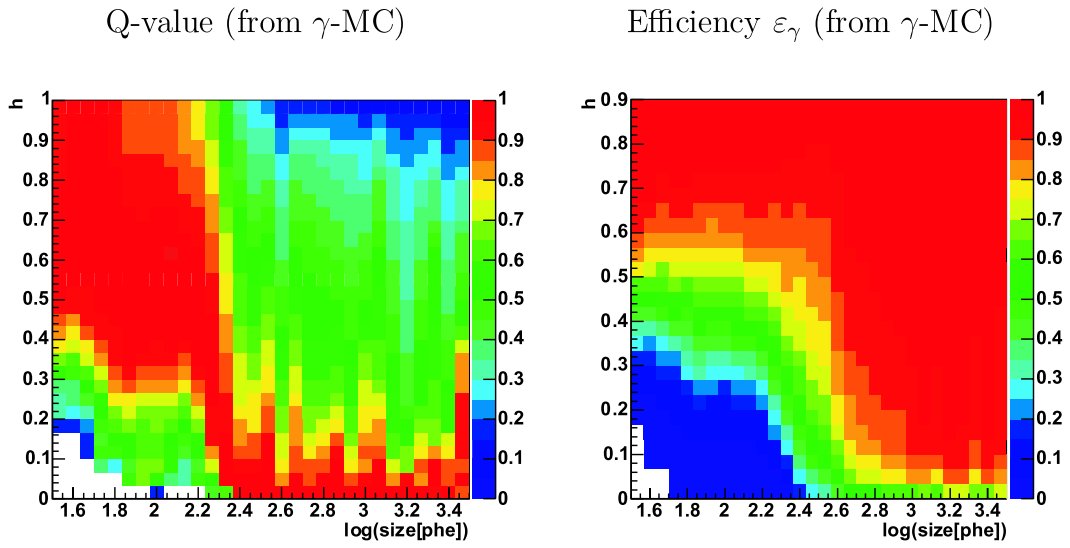
Figure 6.9: Optimization of the g/h-separation using the Q-value. Q is defined as $Q = \frac{\varepsilon_\gamma}{\sqrt{\varepsilon_h}}$ with ε_γ and ε_h being the gamma and hadron acceptances respectively. **Left:** Q *vs* $\log_{10}(size)$. The Q-value distribution in each $\log_{10}(size)$ bin is normalized to the corresponding maximal Q. **Right:** ε_γ *vs* $\log_{10}(size)$

γ-MC
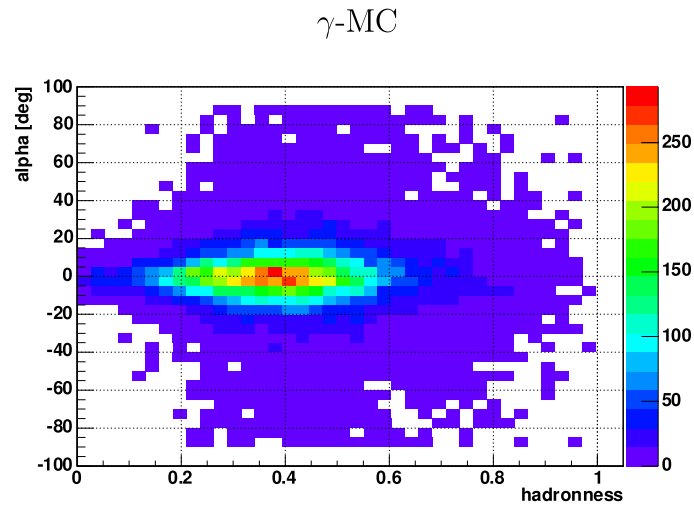
Figure 6.10: γ-MC *alpha vs hadronness* histogram for the bin $70 - 98$ GeV of the estimated energy, which is the first energy bin showing a significant excess. A strong cut in hadronness yields a sharply peaked γ-signal in the alpha plot.

$$S \rightarrow \begin{cases} \sim \frac{N_{ex}}{\sqrt{N_{bg}}} & N_{ex} \ll N_{bg} \\ \sim \sqrt{N_{ex}} & N_{ex} \gg N_{bg} \end{cases} \tag{6.1}$$

Thus, an optimization on $Q = \frac{N_{ex}}{\sqrt{N_{bg}}}$ - using MC data with high γ statistics or ON data from a strong source, preferentially the Crab Nebula - rather than S provides proper cuts for weak sources. In order to be independent of the energy spectrum of the γ-sample under consideration, the optimization on Q must be performed *size*-dependent (see also [Pet97]). Since the MC reflectivity differences amount to not more than about 8 % for the data analyzed in this work (inclusive the PG1553+113 observational data) the *size* dependence is not altered significantly for different data sets.

As can be seen from figure 6.9 a strong cut in hadronness of $h \lesssim 0.1$ is preferable for the whole *size* range $\log_{10}(size) > 2.3$ ($size > 200$ phe).

However, at higher energies corresponding to $size \gtrsim 500$ phe (which roughly corresponds to a region of higher separation quality, compare figure 6.8), a possibly different cut in hadronness will be tested. The optimization is performed on the Crab Nebula ON data itself, since the statistics of the MC gamma test sample may not be sufficient. The cuts obtained in this way will be utilized for the analysis of PG1553+113 data as presented in the next chapter.

For the energy estimation task the RF is trained on the standard image parameters as well as additional parameters including the charge asymmetry parameter $length_q$. In the case of the g/h-separation standard as well as advanced image parameters will be tested. While in the energy estimation additional, advanced image parameters play the role of a correction to the strong energy-*size* dependence, the situation for the g/h-separation is different. In the range of $size < 500$ phe, the traditional *width* and *length* parameters loose their separation power and their 'importance', and new, additional parameters must be carefully checked: Their role in the low *size* range is decisive for the 'hard thresholded' task of accepting or rejecting events.

If not stated otherwise, the following pre-cuts are generally imposed on the ON, OFF and MC data:

- $leakage < 0.1$

- $14^{\circ} < \theta < 30^{\circ}$ (zenith angle range of Crab Nebula ON observations, which covers the zenith angle range of the selected PG1553+113 ON observations)

- 200 phe $< size < 1000$ phe *size* range of significant excess in PG1553+113 ON observations (see next chapter)

The method of cut optimization and the obtained results are given in the following. In addition to the *size*-dependent hadronness cuts supporting cuts in $dist$ and d_μ (see section 4.3.2) are used. As pointed out several times a so-called *alpha*-histogram is used in order to evaluate the number of excess events and its significance. Events having small $dist$ values show a bad 'resolution' in *alpha* (compare figure 5.7). Furthermore, as can be seen in figure 6.11, Cherenkov images from gamma-ray showers show a ring-like distribution of the image's center of gravity: The trigger probability is higher for a location of the center

of gravity in a ring-like area at the edge of the inner camera region. Therefore, strong lower cuts in *dist* are not surprising. Figure 6.12 shows the effect of an upper cut in the 'muon ring center distance' d_μ. The distribution in the *alpha*-histogram becomes flatter.

γ-MC

p-MC
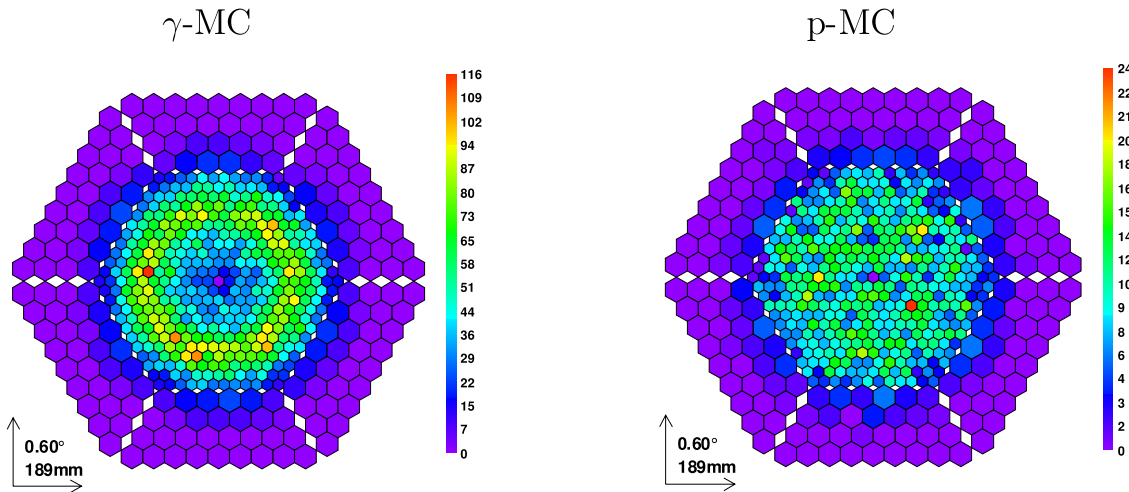
Figure 6.11: Distribution of the center of gravity for gamma (**left**) and proton images (**right**). A standard MC Simulation as introduced in chapter 5 has been used. The ring-like structure in case of the gamma MC has probably the following reasons: The trigger probability at higher *dist* values is smaller since the trigger region covers merely the inner camera. Also the Cherenkov light hump may play a role. Events having small values of *dist* are probably suppressed since the related shower cores are preferentially located high in the atmosphere and far from the telescope, which results in a low Cherenkov photon density.

The additional cuts in *dist* and d_μ thus improve the 'quality' of the *alpha*-distribution. As already mentioned the Random Forest is not trained on the *alpha*-parameter. Thus, RF does not 'know' about the *alpha*-distribution. This is the reason, why in particular the *dist*-cut is used in addition to the hadronness cut, though RF was trained on *dist*. The training of RF on *dist* has its 'right', since the parameters *width* and *length* show, besides of the *size*-dependence, also a *dist*-dependence. It shall be noted that the additional use of cuts in *dist* and d_μ showed no strong influence on the differential energy flux, however, larger Q-values could be achieved. The optimization of g/h-separation cut now proceeds as follows.

1. For the pre-cuts as given above different cuts in the hadronness h are 'scanned' for the following three *size* regions (see also figures 6.8, 6.9 and 6.10):

   - $100 < size < 200 \rightarrow h \lesssim 0.4$
   - $200 < size < 500 \rightarrow h \lesssim 0.1$
   - $size > 500 \rightarrow h \lesssim 0.2$
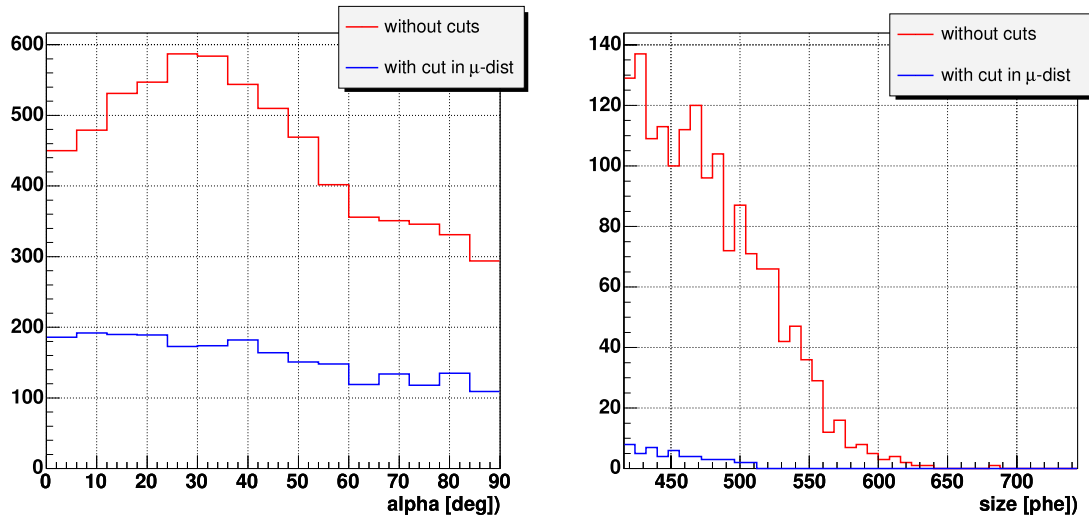
Figure 6.12: Left: *alpha*-histogram for a muon MC as used for the MC calibration without (red curve) and with a cut in $d_\mu < 0.06°$ (blue curve). A hadronness cut of $h < 0.5$ has been additionally applied. **Right:** The *size* distribution as derived from the muon MC. There are no entries for $size > 700$ phe.

'Scanning' means that a generous hadronness cut is taken at first. Then this cut is successively lowered by $\Delta h = 0.01$ and the Q-value is evaluated by means of an alpha histogram. The cut, which provides the maximum Q-value is taken.

2. Different *dist* cuts for the lowest *size* region are 'scanned' (see figure 6.13):

   - $100 < size < 200 \rightarrow dist > 0.6°$

   The above given *dist* cut also provides a maximum Q-value for the higher *size* range, so that it will be taken as static (*size*-independent) cut.

3. Scan *size*-dependent d_μ cut for different *size* ranges (see figure 6.14):

   - $100 < size < 1000 \rightarrow d_\mu < 0.06°$ for $size < 700$ phe

4. Re-scan the hadronness cuts now imposing the *dist* and d_μ-cuts as introduced above (see figure 6.15):

   - $100 < size < 200 \rightarrow h < 0.35$
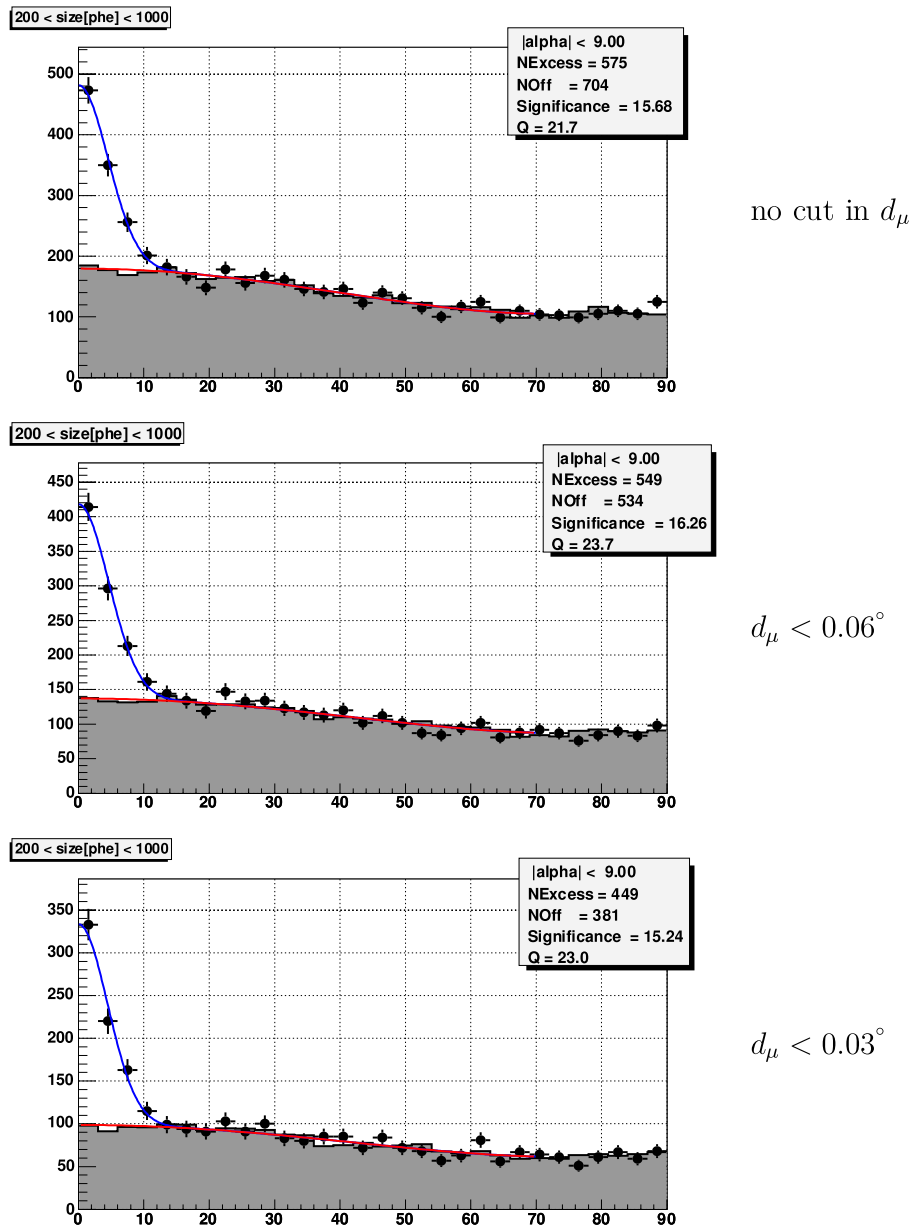   - $200 < size < 500 \rightarrow h < 0.12$
   - $size > 500 \rightarrow h < 0.22$

The final *size*-dependent cuts in hadronness obtained by the above given cut optimization procedure are as follows:

- Precuts and additional cuts:
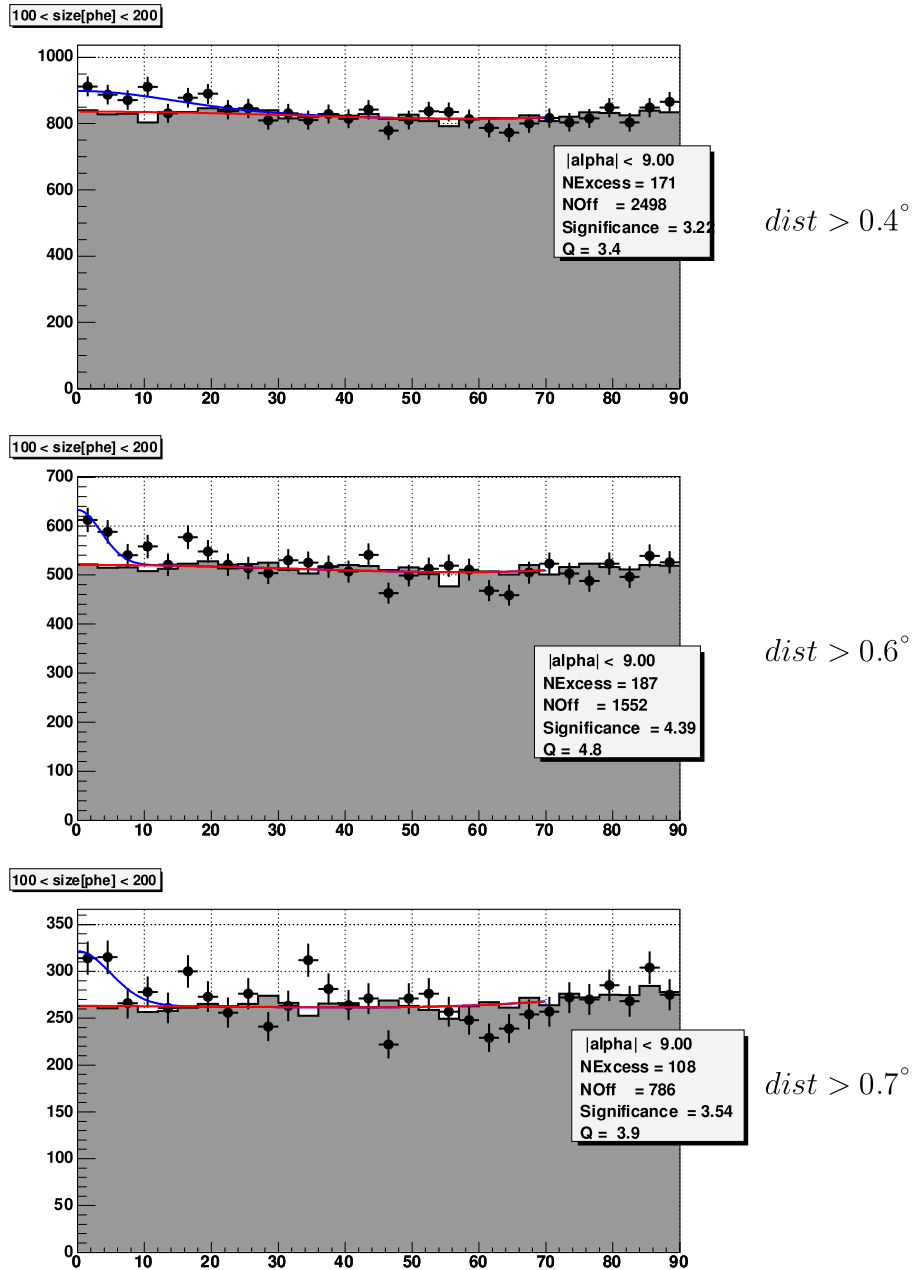
Figure 6.13: Optimization of the *dist* cut for the lowest *size* bin. The red curve shows a fourth-degree polynomial fitted in the range $0^\circ < |alpha| < 70^\circ$ to the scaled OFF-data *alpha*-histogram (see 4.5). This polynomial is used to evaluate the number N_{bg} of background events and the related error in the signal region $|alpha| < 9^\circ$. The blue curve is a Gaussian with zero mean fitted to the ON-data *alpha*-histogram. This fit is not used, it is merely drawn to guide the eye. The number of ON events is calculated directly by summing the histogram bins in the signal region.

Figure 6.14: Optimization of the d_μ cut as described in the text. See figure 6.13 for more explanations.

- – $leakage < 0.1$
- – $dist > 0.6°$
- – $d_\mu < 0.06°$ for $size < 700$ phe

- Standard image parameters:

  - – $100 < size < 200 :\ h < 0.35$ ($Q = 4.80$, $\sigma = 4.39$)
  - – $200 < size < 500 :\ h < 0.12$ ($Q = 14.80$, $\sigma = 10.63$)
  - – $size > 500 :\ h < 0.22$ ($Q = 22.40$, $\sigma = 13.56$)

- Additionally advanced image parameters:

  - – $100 < size < 200 :\ h < 0.28$ ($Q = 2.67$, $\sigma = 2.51$)
  - – $200 < size < 500 :\ h < 0.04$ ($Q = 14.9$, $\sigma = 9.26$)
  - – $500 < size > 1000 :\ h < 0.12$ ($Q = 25.68$, $\sigma = 13.89$)

Figure 6.16 shows the g/h-separation results for the standard image parameters after the above described Q-optimization procedure.

The advanced image parameters introduce a significant improvement for the whole *size* range $200 < size\text{[phe]} < 1000$ (see listing above). However, at the low *size* region there is no significant improvement, but a possible deterioration (see the list with cut values, significances and Q-values given above). Unfortunately it is not possible to check the agreement of the newly introduced parameters completely free of doubt from the figures A.4 to A.7 due to the mentioned difference in the muon content. Figure 6.17 shows the performance of the RF for a training using the set of advanced image parameters and the performance of the classical method of Cuts in Scaled Hillas parameters as introduced in chapter 5, section 5.4.

As already pointed out several times the optimization of the g/h-separation presented in this chapter has the purpose to prepare the analysis of observational data from the BL Lac object PG1553+113, where data with a comparably weak signal from very different observation periods is processed. Therefore, a simple and robust g/h-separation with minimal expected systematic errors is favorable. Furthermore, due to the steepness of the energy spectrum of PG1553+113 a good g/h-separation at comparably low *size* values is needed. Thus the RF g/h-separation based on the standard image parameters will be addressed in the following.

An optimization by means of the Q-value was chosen. It shall be mentioned, that the optimization on the data itself also implies a possible over-optimization for the Crab Nebula analysis (but *not* for the PG1553+113 analysis). The differential energy flux graphs shown in the next section are therefore derived for varying g/h-separation cuts in order to show the reliability of the analysis.
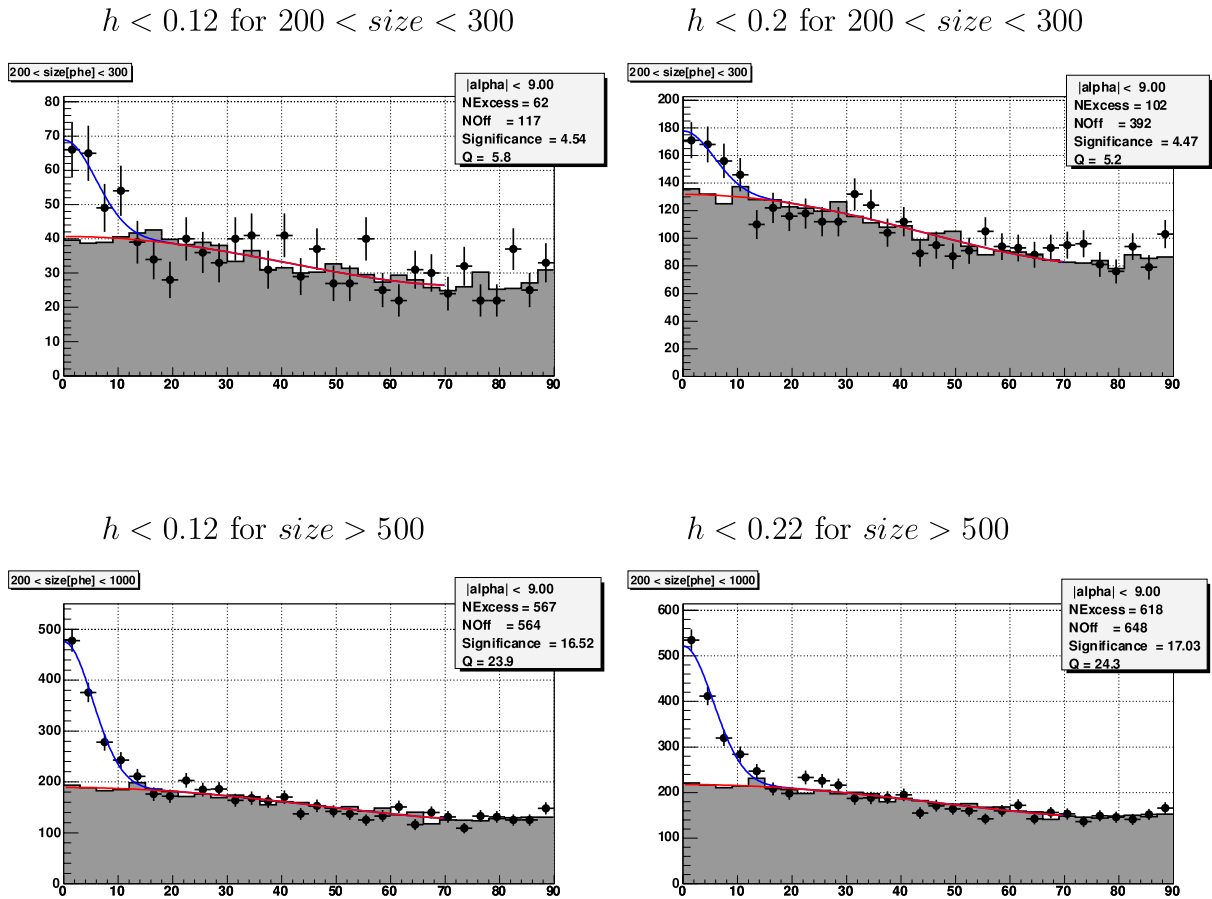
Figure 6.15: Top: Optimization of the hadronness cut in the range $200\,\mathrm{phe} < size < 300\,\mathrm{phe}$. Cuts in $dist$ and d_μ as described in the text have been applied. Even for comparably low $size$ values a strong hadronness cut of $h \lesssim 0.1$ is necessary in order to provide a maximal Q-value. Due to the high OFF data statistics the standard deviation of the Q-value is rather small around $\sigma_Q \approx 0.2$, so that the difference in the Q-values is still significant. An optimization for the size range $200\,\mathrm{phe} < size < 500\,\mathrm{phe}$ leaves these cuts unchanged. **Bottom:** Optimization of the hadronness cut in the range $size > 500\,\mathrm{phe}$. A softer hadronness cut of $h < 0.22$ yields a maximal Q.
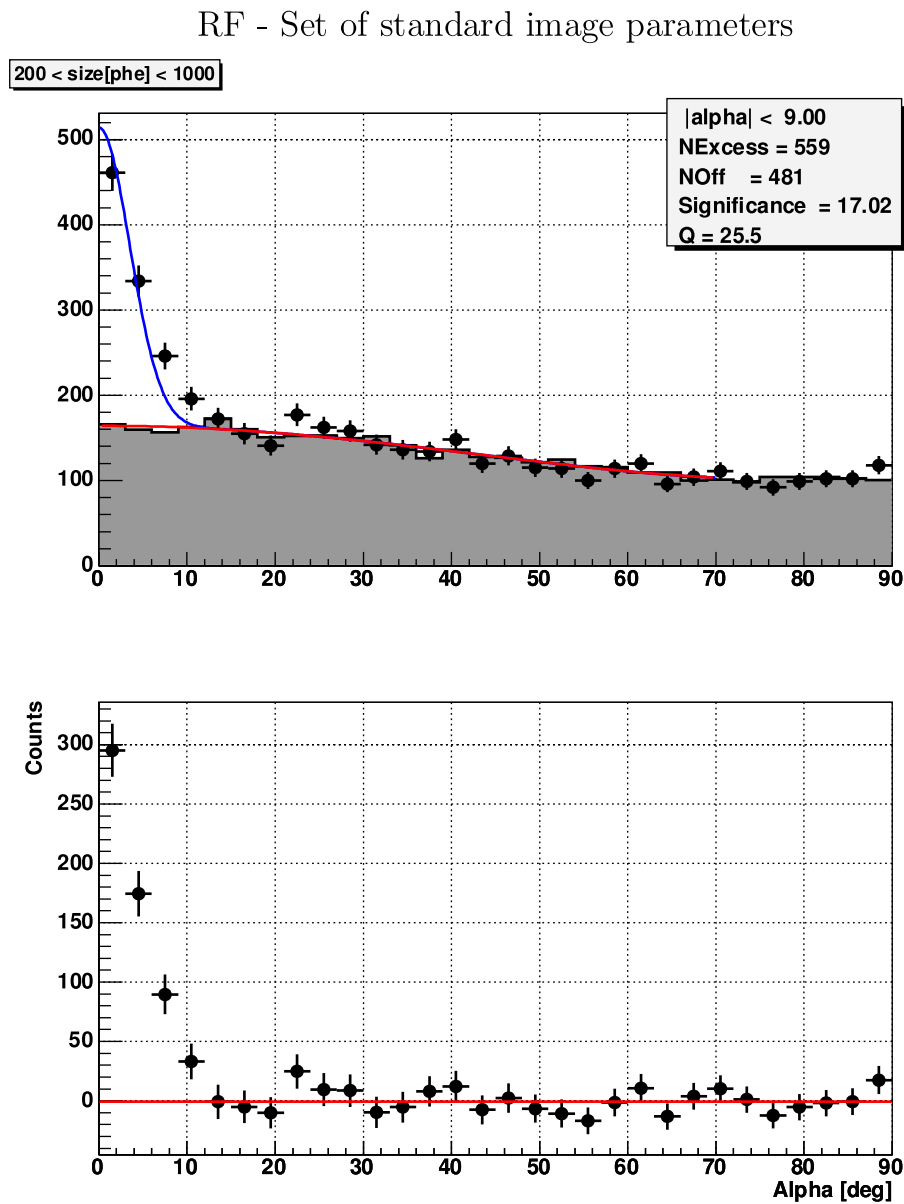
Figure 6.16: Results of Q-value optimization for the set of standard image parameters for the *size*-range $200 < size$[phe] < 1000. The histogram on the bottom shows the *alpha*-histogram after background subtraction.

RF - Set of advanced image parameters

Cuts in 'Scaled Hillas Parameters'
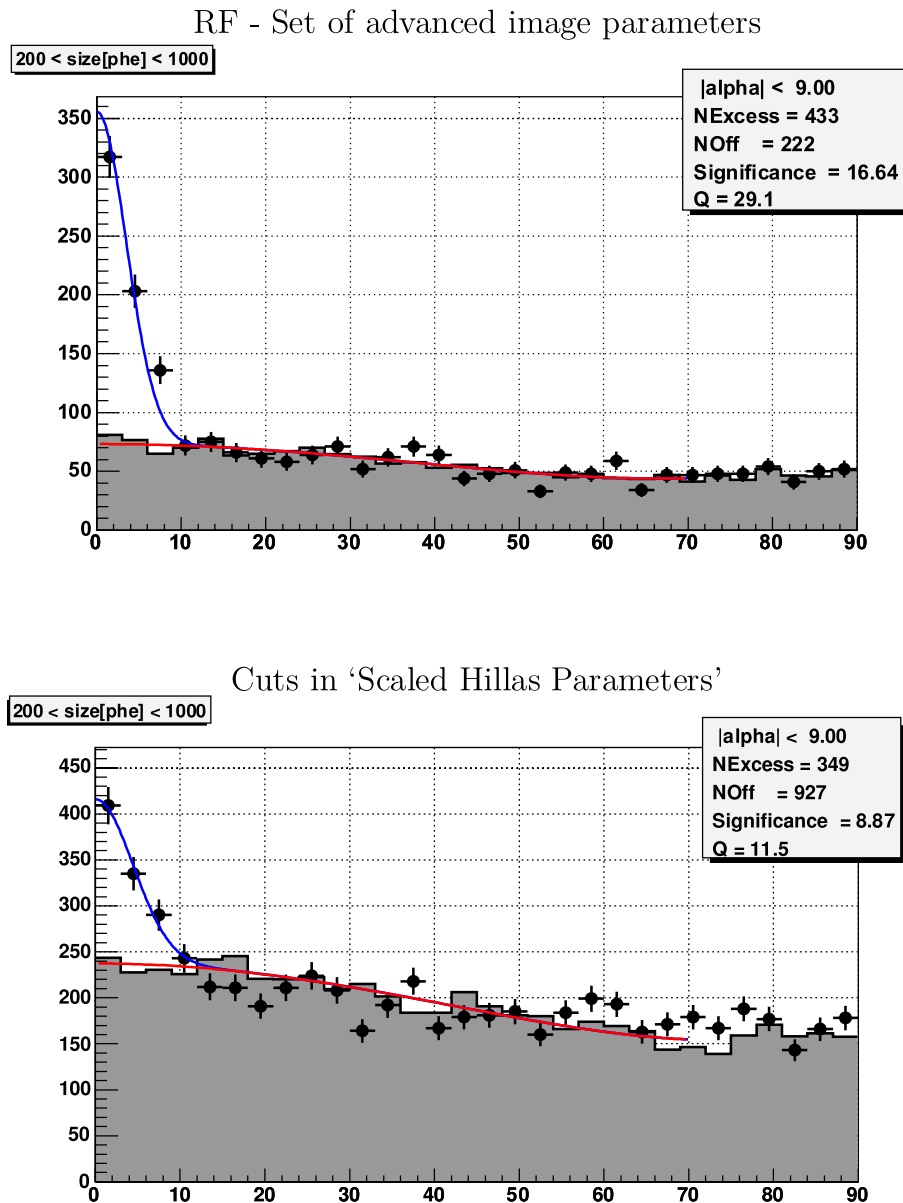
Figure 6.17: Top: Result of Q-value optimization for the Random Forest with the set of advanced image parameters for the *size*-range $200 < size[\text{phe}] < 1000$. **Bottom:** Result of the Q-value optimization of Cuts in Scaled Hillas Parameters as described in section 5.4. Even though the background level seems overestimated in the signal region, a clear advantage of the RF method is visible. The better performance of RF is mainly due to the *size* region below 300 phe.

6.4   Analysis results and discussion

Figure 6.18 shows the data taken during Crab Nebula observations, which is analyzed in the following. The data are recorded on two consecutive days with the modified Julian dates 53269 and 53270.

The so-called *Julian date* (JD) is defined as the number of days since noon on January 1, -4712, i.e., January 1, 4713 BC [Wei06a]. The modified version of the Julian date denoted MJD is obtained by subtracting $2,400,000.5$ days from the Julian date JD.

$$MJD \equiv JD - 2,400,000.5 \tag{6.2}$$

The MJD therefore gives the number of days since midnight on November 17, 1858 [Wei06b].

The effective observation time, which enters into the flux calculation, can be extracted from a histogram of the time differences between succeeding events (see figure 6.18, bottom, and [Wit02a]). If the triggering events are Poisson distributed in time, the time differences Δt should show an exponential distribution $\sim e^{-\lambda \Delta t}$, which can be fitted. The effective observation time T_{eff} is then given as the quotient of the number of the histogram entries N and the constant λ.

In the following g/h-separation settings with different cuts will be used to derive various energy spectra of the Crab Nebula, and their consistency will be checked. The hadronness cut values as derived above will be evaluated and the robustness of the g/h-separation performed by the Random Forest will be checked by varying the hadronness cuts over a wide range.

- g/h-separation A:

$$h < 0.4, \ 100 < size[\text{phe}] < 200 \tag{6.3}$$
$$h < 0.10, \ 200 < size[\text{phe}] < 500 \tag{6.4}$$
$$h < 0.10, \ 500 < size[\text{phe}] < \infty \tag{6.5}$$

- g/h-separation B:

$$h < 0.40, \ 100 < size[\text{phe}] < 200 \tag{6.6}$$
$$h < 0.10, \ 200 < size[\text{phe}] < 500 \tag{6.7}$$
$$h < 0.20, \ 500 < size[\text{phe}] < \infty \tag{6.8}$$

- g/h-separation C:

$$h < 0.40, \ 100 < size[\text{phe}] < 200 \tag{6.9}$$
$$h < 0.20, \ 200 < size[\text{phe}] < 500 \tag{6.10}$$
$$h < 0.20, \ 500 < size[\text{phe}] < \infty \tag{6.11}$$

- g/h-separation D:

$$h < 0.35, \ 100 < size[\text{phe}] < 200 \tag{6.12}$$
$$h < 0.30, \ 200 < size[\text{phe}] < 500 \tag{6.13}$$
$$h < 0.20, \ 500 < size[\text{phe}] < \infty \tag{6.14}$$
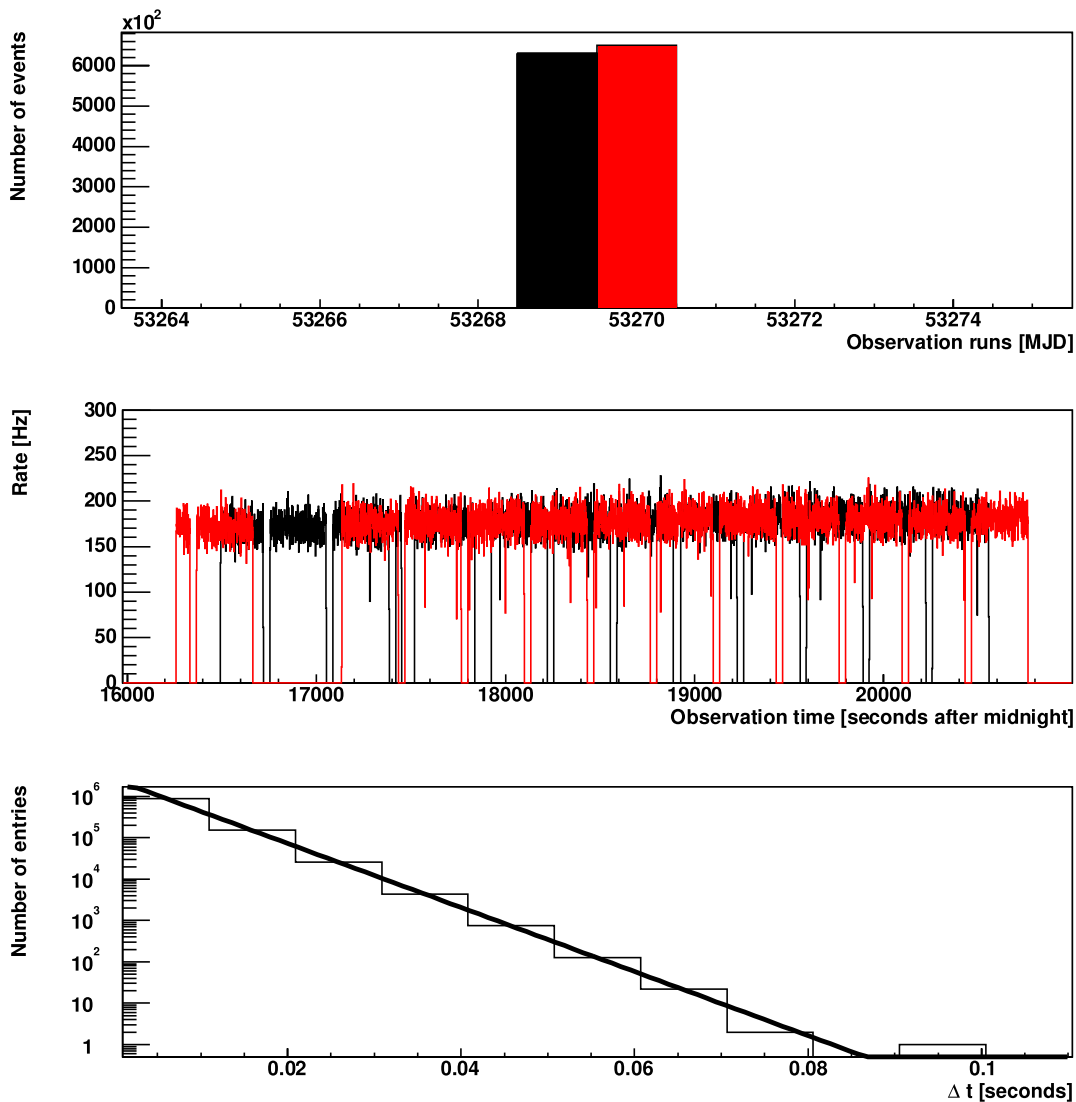
Figure 6.18:   Calculation of effective observation time. **Top:** Histogram of modified Julian date showing the two days, where the Crab Nebula ON data were taken. **Middle:** Histogram of data taking times (the time is given in 'seconds after midnight'). The red and black histograms correspond to the two days of data taking. Since the histogram bin-width is 1 second the rate in Hz can be extracted directly from the bin content. Thus, one can see a rate of ca. $160-210\,\mathrm{Hz}$, with a slight increase with time, likely to be caused by the slightly changing zenith angle. The zero rates in between the data taking are due to calibration and pedestal runs. **Bottom:** Histogram of time differences (time difference between two succeeding events). If the triggering events are Poisson distributed in time, the time differences should show an exponential distribution, which can be used to extract the rate and finally the so-called effective observation time (see [Wit02a]). Here, for the Crab data of the two chosen days a total of $6028\,\mathrm{sec}$ ($1\,\mathrm{h}\,40\,\mathrm{min}$) is calculated for the zenith angle range $14^\circ < \theta < 30^\circ$.

- g/h-separation E:

$$h < 0.35, \ 100 < size[\text{phe}] < 200 \tag{6.15}$$
$$h < 0.12, \ 200 < size[\text{phe}] < 500 \tag{6.16}$$
$$h < 0.22, \ 500 < size[\text{phe}] < \infty \tag{6.17}$$

The differential γ-ray energy spectra are shown in the figures 6.20 and 6.21.

Even for the lowest measurable energies a strong cut in hadronness h is optimal for the following reasons:

- The excess is dominated by $size$ values $size \gtrsim 200$ phe, i.e. $\log_{10}(size[\text{phe}]) \gtrsim 2.3$, where the Q-value is maximized for strong cuts in hadronness (see figure 6.9).

- The $|alpha|$-distribution becomes significantly narrower for a hard g/h-separation cut (see figure 6.10). With a looser hadronness cut the width of the γ-signal in the α-plot could be wider and it would be more difficult to determine the background in the signal region.
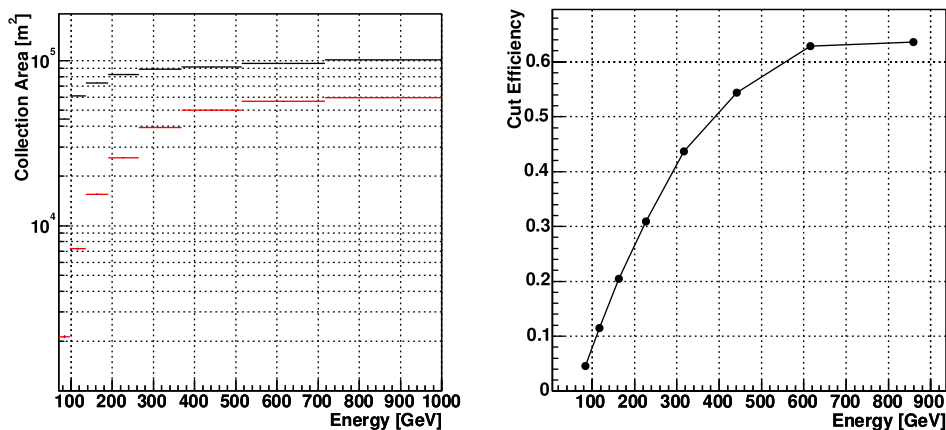
Figure 6.19: Plots documenting the strong g/h-separation A for image parameter set 1. **Left:** Effective collection area as a function of energy. The zenith angle range covers an interval $14° < \theta < 30°$, similar to that of the chosen PG1553+113 observations. The black and red histograms show the collection area before and after the g/h-separation cuts, respectively. The very small and negligible statistical error of the collection area is indicated by a vertical error bar, which is for some bins smaller than the thickness of the line. **Right:** Gamma cut efficiency as a function of energy.

However, there is a problem in using strong cuts with low γ cut efficiency: The MC simulation must be well adjusted, so that the effective collection areas are precisely known and no systematic errors enter into the flux calculation. The situation can be checked by determining energy spectra for both soft and hard cuts, which is the main aim of this

chapter. Figures 6.20 and 6.21 show Crab Nebula VHE γ-ray spectra for the different g/h-separations.

There is good agreement between the differential γ-ray fluxes. A comparison with results of an independent analysis yields a good agreement within the statistical errors (see figure 6.1). However the first energy bin shows a significant tendency, which indicates that the systematical error here is in the order of the statistical error.

The Crab flux as measured by the HEGRA experiment (red, dashed line in figures 6.20 and 6.21) can be well confirmed. This differential flux given in [A$^+$04] as

$$\frac{dF}{dE} = (2.83 \pm 0.04_{stat} \pm 0.6_{sys}) \cdot 10^{-11} \cdot \left(\frac{E}{1\,\text{TeV}} \right)^{-2.62 \pm 0.02_{stat} \pm 0.05_{sys}} \quad \text{photons}\,\text{cm}^{-2}\,\text{s}^{-1}\,\text{TeV}^{-1}$$

was measured for energies $E > 500\,\text{GeV}$. The red, dashed line as drawn in figures 6.20 and 6.21 is an extrapolation. The differential energy flux for energies $E > 300\,\text{GeV}$ determined in this work is

$$\frac{dF}{dE} = (2.93 \pm 0.2_{stat} \pm 0.10_{sys}) \cdot 10^{-11} \cdot \left(\frac{E}{1\,\text{TeV}} \right)^{-2.56 \pm 0.15_{stat} \pm 0.02_{sys}} \quad \text{photons}\,\text{cm}^{-2}\,\text{s}^{-1}\,\text{TeV}^{-1}$$

The systematic error was estimated from the different g/h-separations. The differential flux below $300\,GeV$ can be crosschecked using the spectrum as given in [WM05]. Also here, a good agreement is evident (see figure 7.7). The lowest significant point (the flux value below $100\,\text{GeV}$) shows significances of 3σ for the strongest g/h-separation up to nearly 5σ for the softer cuts.
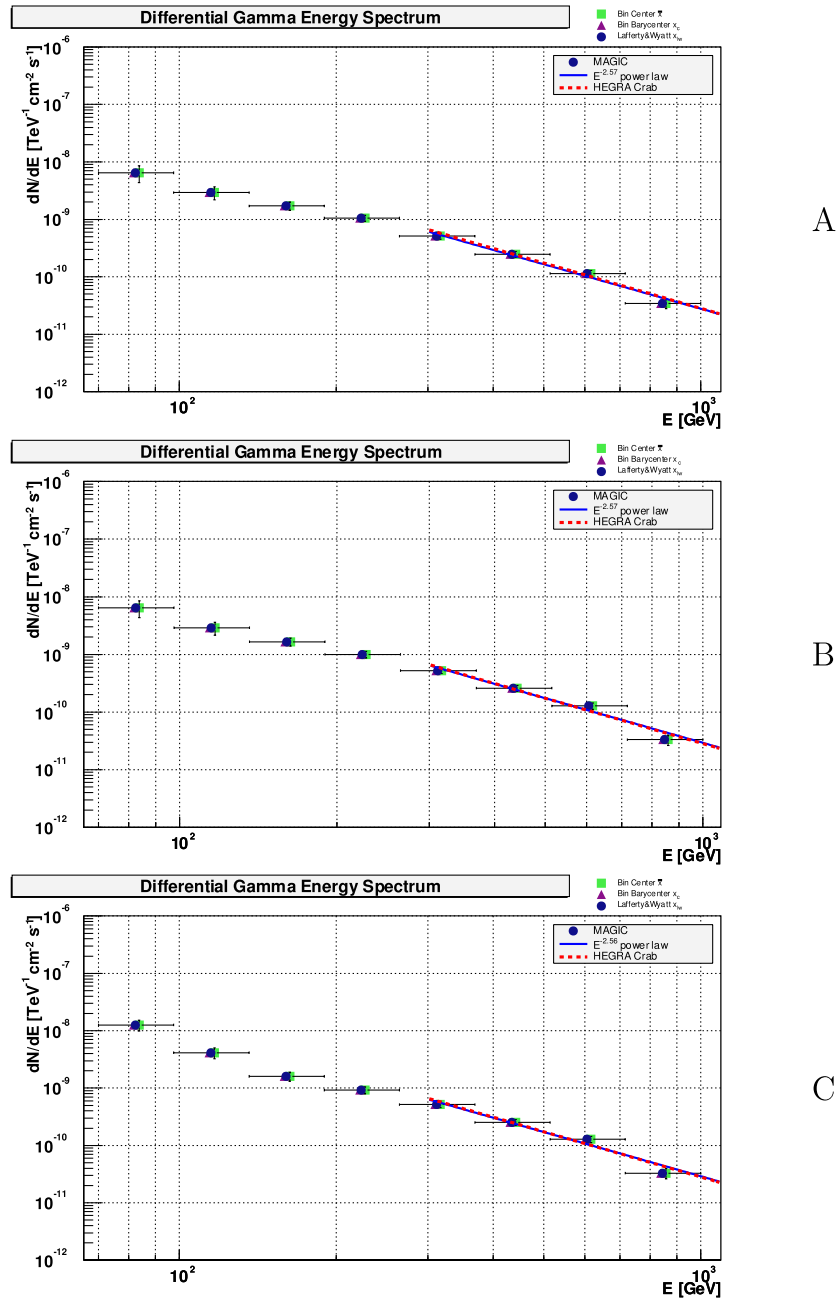
Figure 6.20: Differential γ-ray flux calculated from Crab Nebula ON data using the image parameter set 1 (see text) and g/h-separations A, B and C corresponding to the plots on top, middle and bottom respectively. The differences between A, B and C lie within the statistical errors. The dashed red line denotes the Crab Nebula VHE flux as measured with the HEGRA experiment, extrapolated down to 300 GeV. The errors of the flux levels are typically large at very high energies (low flux level) and very low energies (high flux level but less efficient g/h-separation).
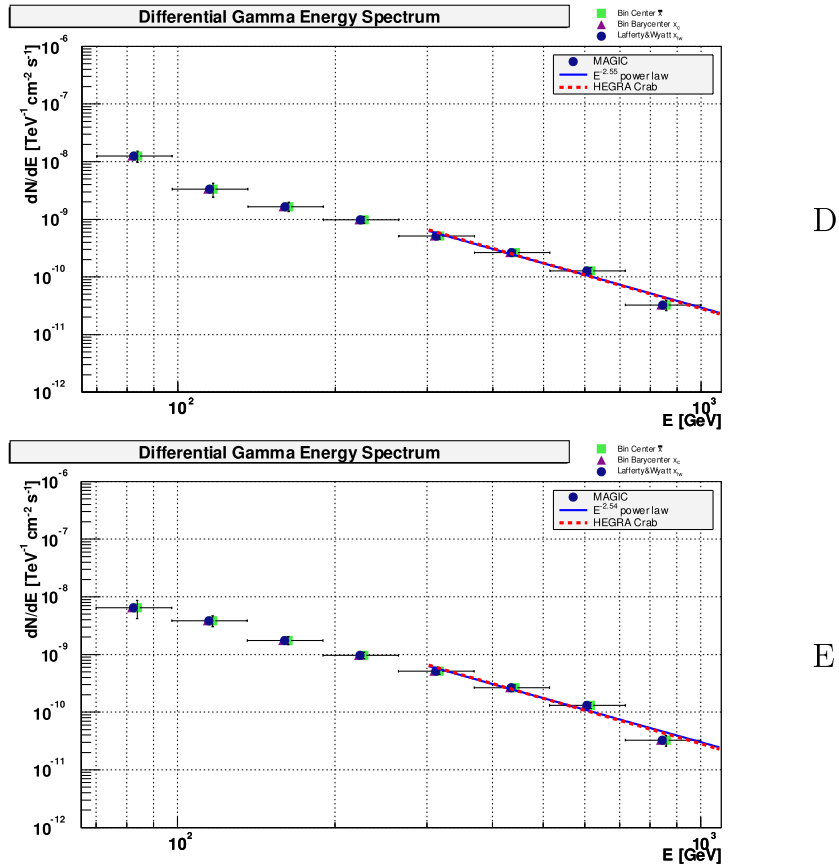
Figure 6.21: Differential γ-ray flux calculated from Crab Nebula ON data using the standard image parameters and g/h-separations D and E corresponding to the plots on top and bottom respectively. g/h-separation E is that one which yields maximum Q-values and will be taken for the source analysis presented in the next chapter. For further explanations, see figure 6.20. The measured spectral index in the energy range $E > 300\,\text{GeV}$ varies as $\Gamma = 2.56 \pm 0.15_{stat} \pm 0.02_{sys}$ for the spectra A-E. See also figure 7.7 in the next chapter for a comparison with the differential energy flux of PG1553+113 and with the results obtained with the standard MAGIC data analysis [WM05].

Chapter 7

# Analysis of the VHE γ-emission from the BL Lac PG1553+113

7.1 Overview

The AGN PG1553+113 was discovered in the Palomar-Green survey of UV-excess stellar objects [GSL86] (hence the catalog name 'PG' followed by the usual naming convention using right ascension and declination). It was classified as BL Lac object due to the absence of spectral emission lines [MG83, FT90] and the strong optical variability of $m_p = 13.2 - 15.0$ [MG83], with m_p being the so-called apparent magnitude (a logarithmic scale of luminosity running from small values of magnitude for the brightest stars to large values for the faintest, see e.g. [Per03]). The featureless spectrum is typical for blazars since the non-thermal emission dominates over thermal radiation due to beaming effects [ST06]. Also typical is the lack of a UV-excess or 'blue bump' (originating from the accretion disk thermal radiation) [LMKF$^+$99]. The classification as High Peaked BL Lac (HBL) is based on the synchrotron peak position of the SED [GAM95]. The SED was measured in several multiwavelength campaigns, the X-ray region by the experiments BeppoSAX [DSG05] and XMM Newton [P$^+$05]. There exists an upper limit in the gamma-ray region from EGRET observations [F$^+$94].

The redshift of the source is still unknown. Originally, a redshift $z = 0.36$ was attributed to PG1553+113 [MG83], but this measurement was based on a spurious emission line, and later on revised [FT90]. Emission lines could principally be created by fluorescence radiation from clouds surrounding the central black hole. Absorption may occur due to gas in the halo of foreground galaxies as it happens for some high-redshift quasars or due to source intrinsic gas. The first effect is better measurable when the source is in a high activity state, the latter when the source is in a low state [STF05].

However, no emission or absorption lines could be detected and therefore only estimates for the redshift of PG1553+113 can be given:

- In a Hubble Space Telescope (HST) survey of 110 AGN the host galaxies of all objects with redshift $z < 0.3$ could be resolved (see figure 7.1). Since the host galaxy of PG1553 is not resolved a lower limit of $z > 0.3$ can be given as estimate (see [CFK$^+$02]).

- Assuming a typical host galaxy with absolute magnitude $M_H = -22.9$ [STF05] and taking into account a general relation between the redshift z and the apparent magnitude m_H a lower limit of $z > 0.78$ is found after imposing an upper limit on m_H (host galaxy unresolved in the mentioned HST survey) [STF05].

- The *nucleus to host flux ratio* ρ is a function of the nucleus and host galaxy's apparent magnitudes and therefore of z. Since the so-called equivalent width (EW) of a spectral absorption line also depends on ρ and therefore on the redshift z, an upper limit of the EW deduced from the non-detection of any absorption lines can be used to derive a lower limit for z (see [ST06] for the definition of EW and more details concerning this method). In this way a lower limit of $z > 0.09$ could be derived from ESO-VLT optical spectroscopy observations of PG1553+113 [ST06].
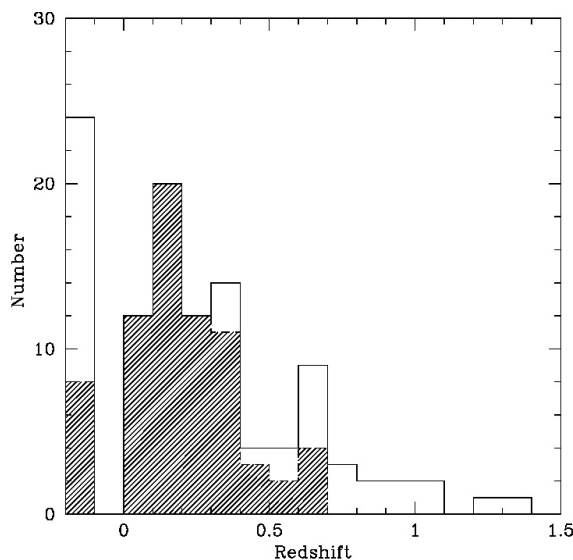

Figure 7.1: Redshift histogram for BL Lac type AGN observed in a HST survey of 110 objects[Urr00]. The cross hatched area indicates the BL Lac objects with resolved host galaxy. For redshifts $z > 0.5$ only 6 out of 22 host galaxies are resolved. The first bin ($z < 0$) depicts all objects with unknown redshift, here only one third have resolved host galaxies. From this diagram one can conclude a lower redshift limit of $z > 0.3$ for PG1553+113 related to the bin upper edge of the last redshift bin, which shows only objects with resolved host galaxy. Picture taken from [Urr00].

7.2   Analysis results

Figures 7.2 and 7.3 show some details concerning the data taken in the PG1553+113 observations by the MAGIC telescope in 2005 and 2006. Runs with strongly varying and low rates have been excluded (see figure 7.4). The observation time after the exclusion of
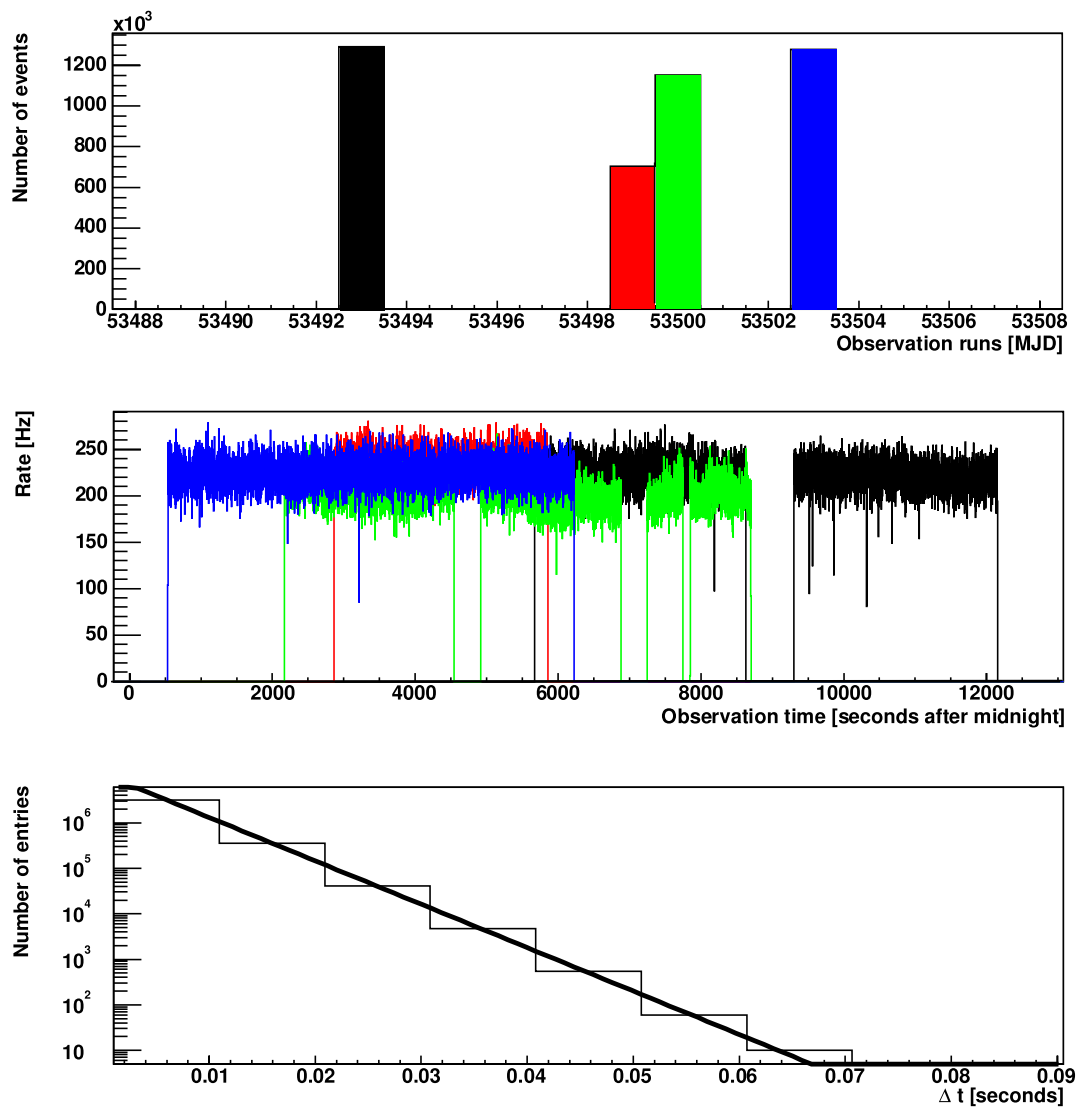
Figure 7.2: Overview of PG1553+113 observations in 2005. **Top:** Histogram of modified Julian date showing the days, where the PG1553+113 ON data were taken. **Middle:** Histogram of data taking times (the time is given in 'seconds after midnight'). The differently colored histograms correspond to different days of data taking. Since the histogram bin-width is 1 second the rate in Hz can be extracted directly from the bin content. Thus, one can see a rather stable rate with minimum and maximum values in the range of about $180 - 250$ Hz. **Bottom:** Histogram of time differences (time difference between two succeeding events). For further details see figure 6.18
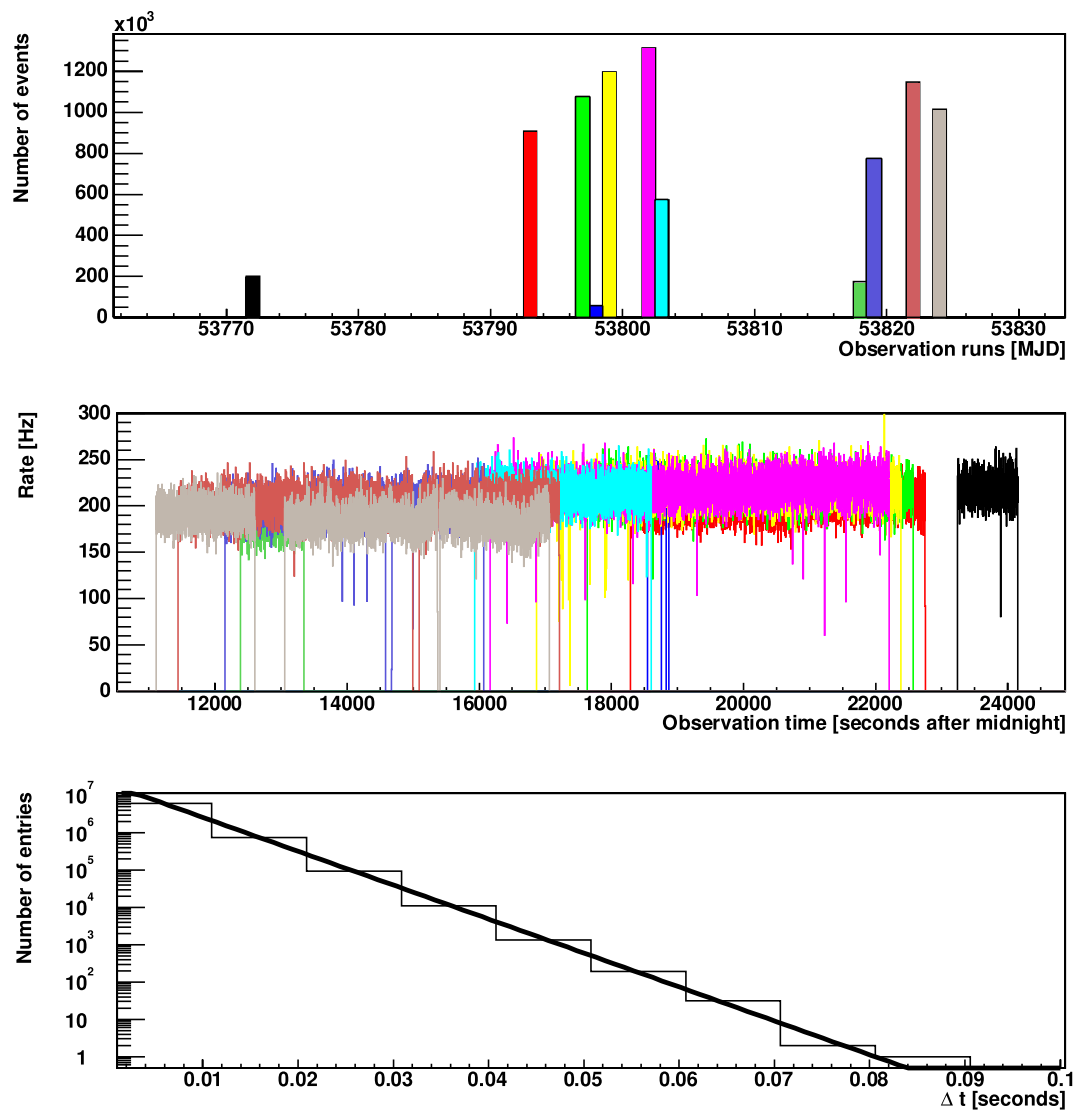
Figure 7.3: Overview of PG1553+113 observations in 2006. The observation time after exclusion of runs with varying and low rates is about 11 h 19 min. For explanations, see figures 7.2 and 6.18.
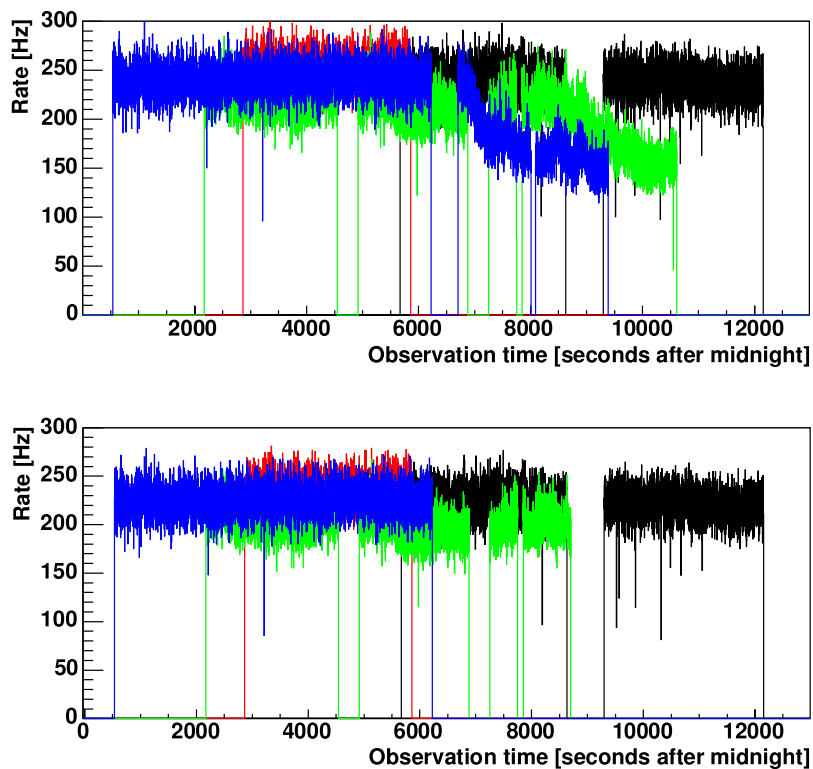
Figure 7.4: Top: All PG1553+113 data runs in 2005. The pure data taking time excluding pedestal and calibration runs amounts to 6 h 52 min. **Bottom:** Selected PG1553+113 data runs in 2005. Runs and fractions of runs with stable rates were selected. The pure data taking time after this selection amounts to 5 h 37 min. The data runs in 2006 were as well selected by checking the rate stability.

'bad runs' amounts to 5 h 37 min in 2005 and 11 h 19 min in 2006, thus data corresponding to a total observation time of 16 h 55 min is analyzed in the following.

In analogy to the Crab Nebula analysis as presented in chapter 6 reflectivity and PSF of the MC simulation were adjusted by means of a muon calibration. Figures A.9, A.8 and A.10 (see appendix A) show the graphs for muon *arc width* and *size* vs *arc radius* distributions for PG1553+113 2005 ON data and the 'best fit' muon MC simulation.

The g/h-separation optimized on the Q-value with cuts in *dist*, d_μ and hadronness as determined in the Crab Nebula analysis of last chapter has been applied. Figure 7.5 shows the *alpha* histograms for the different data sets of 2005 and 2006 as well as for the combined data. Table 7.1 shows the corresponding numbers of excess and background events as well as the significances. A clear excess $\gtrsim 5\sigma$ is present in all data sets. The significance as shown in the *alpha* histograms was calculated according to equation 4.50 (denoted in table 7.1 by the subscript *LiMa5*). The significances evaluated by means of equation 4.49 with the effective values given in equation 4.52 are larger (denoted in table 7.1 by the subscript *LiMa17*), they are given here for comparison with the results obtained in [A$^+$06c].

| Year	N_{ex}	N_{bg}	$\sigma_{N_{bg}}$	S_{LiMa5}	S_{LiMa17}
2005	222	1528	14.5	5.01	5.15
2006	526	3564	33.7	7.27	7.45
2005+2006	747	5092	48.0	8.27	8.42

Table 7.1: Number of excess and background events as well as significances for the different sets of PG1553+113 observational data. A range of $|alpha| < 9^\circ$ has been used for the calculation of the number of background and excess events.

The false source histogram for PG1553+113 is given in figure 7.6. The signal at the center of the histogram is clearly visible. For comparison the false source histogram for the Crab Nebula is given for the same *size* range in order to demonstrate the expected resolution.

The spectrum of PG1553+113 as determined from the observations in 2005 and 2006, together with the overlaid spectrum derived from Crab Nebula observations in 2004 are depicted in figure 7.7. The analysis of data taken during Crab Nebula observations as described in the last chapter revealed a systematic error of the flux level at energies below 100 GeV, which is in the order of the statistical error. At these low cut efficiencies the collection areas are expected to be unreliable. Thus, the first data point at 90 GeV was omitted in the power law fit. The flux level of PG1553+113 ranges from $\sim 4.6\,\%$ Crab flux for the last energy bin ($E = 500$ GeV) up to nearly the Crab level for the lowest energy bin at $E = 90$ GeV. However, a detection below 90 GeV (in comparison to the data point at $E = 80$ GeV in the Crab Nebula energy spectrum) was not possible, which may indicate a spectral break.

Figure 7.8 shows the *alpha*-histogram of the MAGIC data analysis presented in [A$^+$06c]
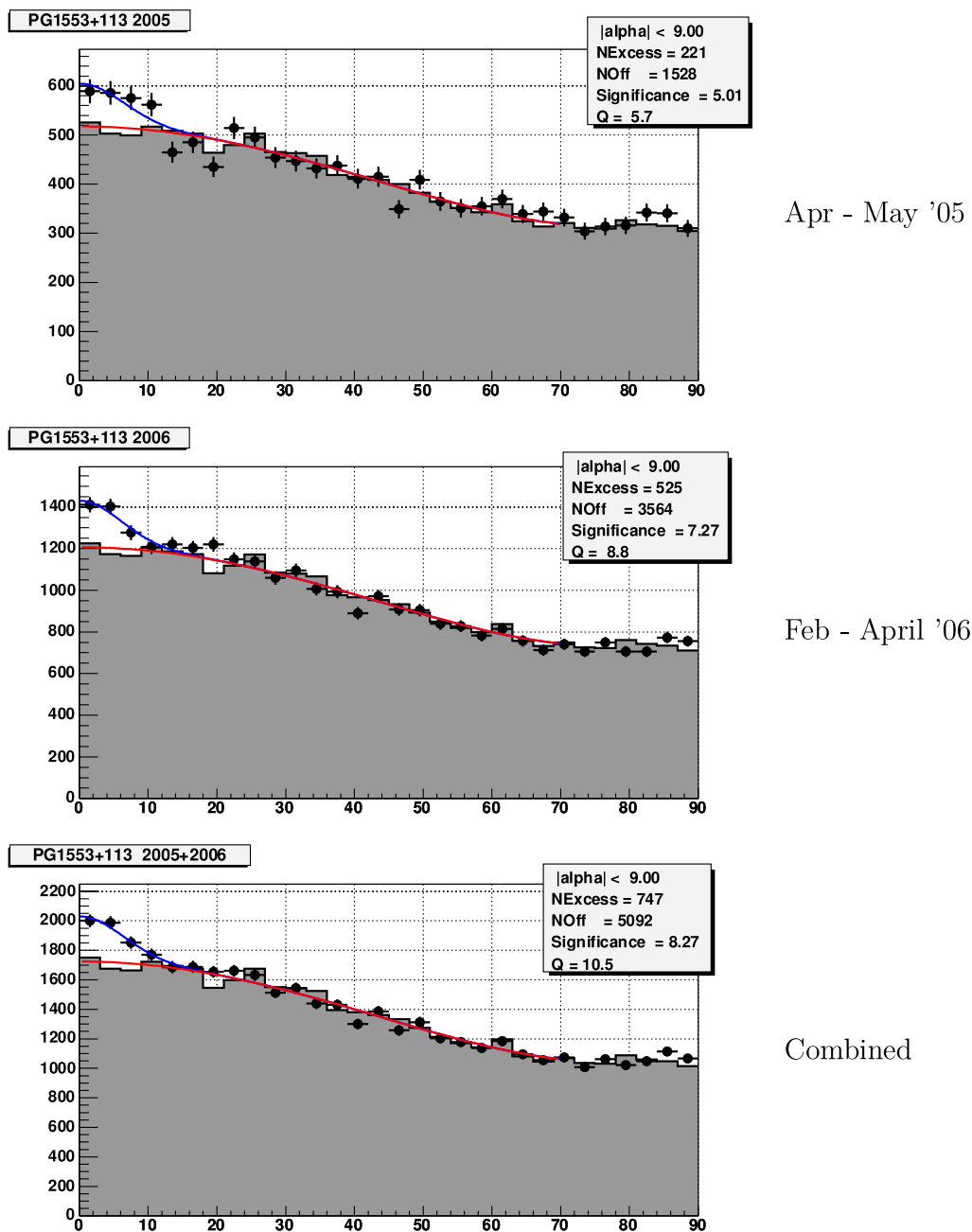
Figure 7.5: Alpha plots for different sets of observational data taken on PG1553+113.
Top: Data taken in April and May 2005. **Middle:** Data taken in February to April 2006.
Bottom: Combined 2005 and 2006 data sets. The g/h-separation cuts as introduced for
the Crab analysis (previous chapter) have been used. The curved shape of the background
level can be attributed to different reconstruction efficiencies for different positions of the
showers in the camera (see figure 3 in [Wit02b]). The wavy background level is also visible
in the *alpha* histograms shown in the Crab Nebula analysis of the last chapter, if one takes
into account the strong gamma signal of the Crab Nebula, which is dominating the shape
of the *alpha* histogram.
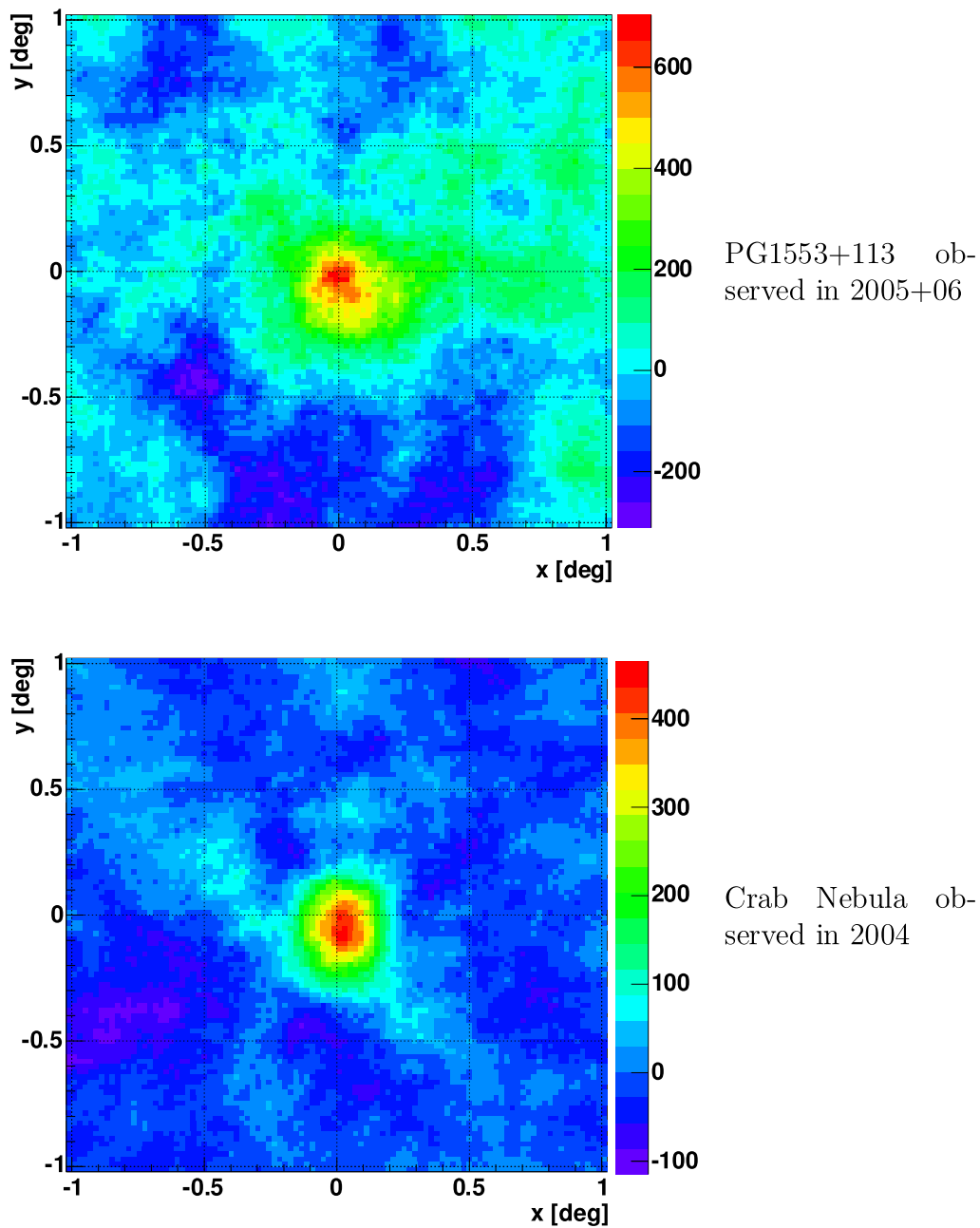
Figure 7.6: False source histograms showing the number of excess events (see also section 4.5). **Top:** False source histogram for the combined PG1153+113 data set **Bottom:** False source histogram for the Crab Nebula data set analyzed in the last chapter. The size range was restricted to $200 < size[\text{phe}] < 1000$, so that the resolution is about the same as for the PG1553+113 false source histogram. A slight mispointing of about $0.05°$ (half a pixel) in both x and y directions becomes visible, which is of the order of the PSF $(0.04°)$. It shall be mentioned that a pointing correction did not show an influence on the g/h-separation cuts and the energy spectra presented here.

for all data. The analysis of $18.8\,h$ PG1553+113 observations with zenith angle $0^{°} < \theta < 30^{°}$ yields an excess of 1032 ± 119 at a background level of 8730 ± 66 which results in a significance of $S = 8.8$. A selection of the same zenith angle range as in this work was done, but the data presented in this chapter corresponds to a total observation time of 16 h 55 min, which may be due to different selection criteria. As can be seen be comparing figures 7.5 and 7.8 following differences are apparent for the data analysis in [A⁺06c]:
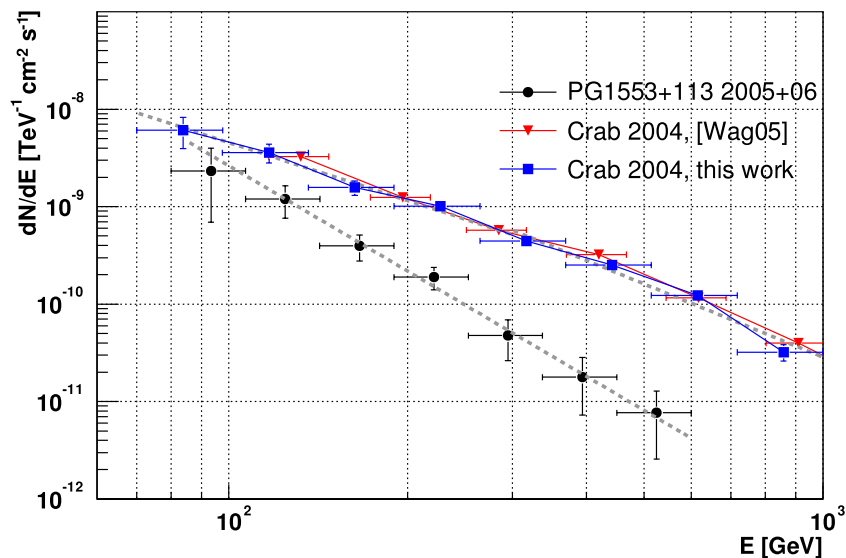
Figure 7.7: The differential γ-ray energy spectrum of PG1553 as determined from data taken in 2005 and 2006 (see figure 7.5). The flux level of PG1553+113 reaches about 60 % Crab flux at 100 GeV and decreases to about 4.6 % Crab flux at 500 GeV. The photon index of the Crab energy spectrum drops from $\Gamma = 2.5 \pm 0.15$ to $\Gamma = 2.0 \pm 0.1$ when fitting a power law separately for the energy ranges $300\,\text{GeV} < E < 1000\,\text{GeV}$ and $70\,\text{GeV} < E < 300\,\text{GeV}$, respectively.

- *Background level:* In the range $alpha < 45^{°}$ only one data point lies below the fitted background level (dashed line), which indicates a (slightly) underestimated background level.

- *Position of alpha-peak:* The peak of the *alpha*-histogram is located in the second bin corresponding to $3^{°} < alpha < 6^{°}$. The first bin shows an atypically small count number, which nevertheless could be just a fluctuation. However, such a behaviour could not be seen in the analysis presented in this thesis, which rules out a possible mispointing as reason.

- *Width of alpha-peak:* The *alpha*-peak shows a rather broad shape extending beyond $12^{°}$. For the presented analysis the signal region of the ON data *alpha*-histogram corresponds to the first 3-4 bins $(9^{°} - 12^{°})$.

Taking into account the just mentioned differences together with the also differing observation times the significances obtained in the analysis of this work are well comparable.
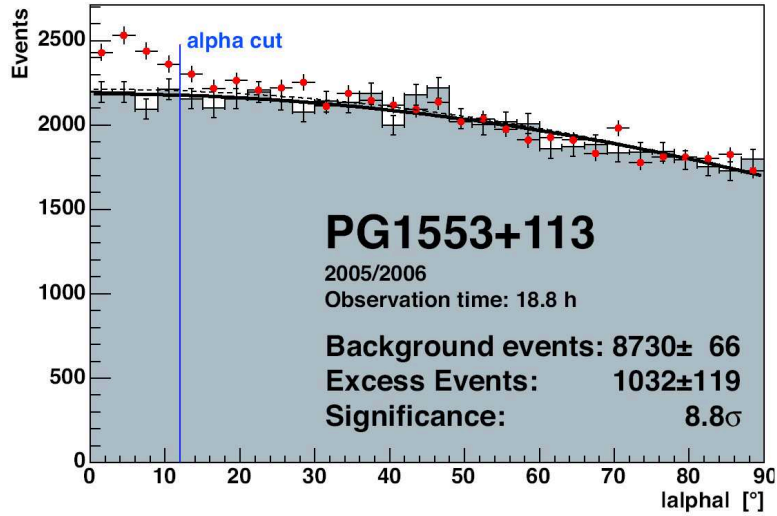
Figure 7.8: *alpha*-histogram derived from observations by the MAGIC telescope in 2005 and 2006 [A$^+$06c] (compare with figure 7.5). For the method of calculating errors and significances see section 4.5, which is also valid for the standard MAGIC analysis.
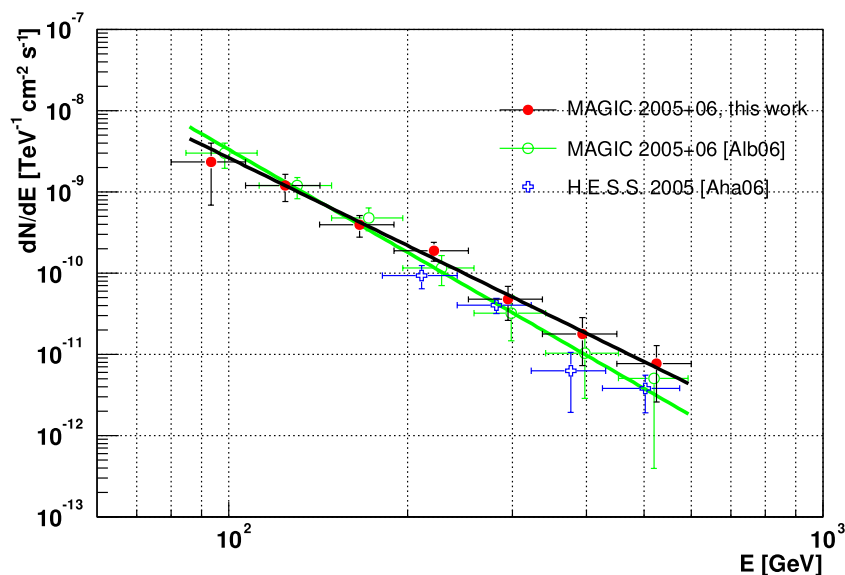
Figure 7.9 shows differential energy spectra of the VHE gamma-emission from PG1553 +113 as observed by the H.E.S.S. and MAGIC experiments. The MAGIC data set of the second analysis [A$^+$06c] roughly corresponds to the data which is analyzed in this work. The flux levels and spectral indices of the two MAGIC data analyses are in agreement within their statistical error. However at the 'high energy end' of the differential energy flux there is a tendency to higher flux values. Yet, here the significance of the detected signal decreases resulting in comparably big errors of the flux level. The differential energy flux of PG1553+113 as measured in this work is (in concordance with [A$^+$06c])

$$\frac{dF}{dE} = (0.187 \pm 0.036_{stat}) \cdot 10^{-9} \cdot \left(\frac{E}{200 \text{ GeV}}\right)^{-3.6 \pm 0.3_{stat}} \text{ photons cm}^{-2} \text{ s}^{-1} \text{ TeV}^{-1}$$

The integral flux F above 200 GeV and the photon index Γ for the different analyses as given in [A$^+$06b], [A$^+$06c] and in this work are as follows:

H.E.S.S. [A$^+$05]: $F_{>200 \text{ GeV}} = (0.48 \pm 0.1_{stat} \pm 0.1_{syst}) \cdot \frac{10^{-12}}{\text{cm}^2 \text{ s}}$ $\Gamma = 4.0 \pm 0.6_{stat}$
MAGIC [A$^+$06c]: $F_{>200 \text{ GeV}} = (1.0 \pm 0.4_{stat}) \cdot \frac{10^{-12}}{\text{cm}^2 \text{ s}}$ $\Gamma = 4.2 \pm 0.3_{stat}$
MAGIC this work: $F_{>200 \text{ GeV}} = (1.7 \pm 0.3_{stat}) \cdot \frac{10^{-12}}{\text{cm}^2 \text{ s}}$ $\Gamma = 3.6 \pm 0.3_{stat}$

In this work no further correction of the energy spectrum was applied (see also comments concerning the reconstructed energy spectrum at the end of section 5.5.1). However

Figure 7.9: The differential γ-ray energy spectrum of PG1553+113 as determined by H.E.S.S. in 2005 and MAGIC in 2005 and 2006 (see figure 7.8). The last data point in the flux of the independent MAGIC data analysis [A$^+$06c] is not a measurement but an upper limit.

in the analysis presented in [A$^+$06c] a correction for the energy resolution was used. This may explain the smaller integral flux value found in [A$^+$06c], which nevertheless is within statistical errors.

The excess rate for the two data sets of 2005 and 2006 is rather compatible (see figure 7.5 and table 7.1, where a ratio of the excess per observation time of about 0.8 can be extracted). It is therefore reasonable to add these data sets and determine the differential energy flux as shown. A separate treatment of the data sets reveals a higher integral flux ($F_{>200\,\text{GeV}}$) for the 2005 data set, though not with good significance (the low statistics actually does not allow a thorough determination of the spectra for the separated data sets, especially for energies $E > 200$ GeV, and thus the integral flux above 200 GeV is quite uncertain).

Furthermore, it was not reasonable to produce flux estimates for smaller time intervals, since the errors were too large to draw definite conclusions from such a light curve.

The analysis of Crab Nebula observational data indicated a systematic error of the flux level right below 100 GeV likely to be caused due to the extremely low cut efficiencies. Therefore, the further analysis (in particular the implications of the energy spectrum's steepness on the EBL as given below) will disregard the related data point.

7.3   Discussion of the energy spectrum

A detailed discussion concerning the extinction of γ-rays in the metagalactic radiation field was already given in section 1.5. In preparation of the further analysis a simple SED model of blazar spectra using the approach of Krawczynski [K$^+$03], which is a one-zone Synchrotron Self Compton (SSC) simulation, is introduced. A deeper insight into the theoretical background of analogous models is given elsewhere [K$^+$99, IT96], here a comparably compact introduction is given.

Within a simple SSC model approach the γ-ray emission is assumed to be produced inside a spherical blob of radius R (typical dimension $R \approx 10^{16}$ cm) containing a non-thermal, homogeneous and isotropic electron population moving with relativistic speed towards the observer (for blazar type AGN the jet axis is roughly identical to the line of sight).

The time evolution of the electron number density $n_e(\gamma, t)$ is described by following differential equation [KTK02].

$$\frac{\partial n_e(\gamma, t)}{\partial t} + \frac{n_e(\gamma, t)}{t_{e,esc}} \;\; = \;\; -\frac{\partial}{\partial \gamma} \left[(\dot{\gamma}_{syn} + \dot{\gamma}_{ssc}) \cdot n_e(\gamma, t) \right] + \Im_{e,inj}(\gamma, t) \tag{7.1}$$

$$\Im_{e,inj}(\gamma, t) \;\; = \;\; q_e \gamma^{-p} e^{-\frac{\gamma}{\gamma_{max}}} \, , \; \gamma > \gamma_{min} \tag{7.2}$$

Here, $t_{e,esc}$ denotes the effective escape time of the electrons. γ is the electron Lorentz factor and $\dot{\gamma}_{syn}$ and $\dot{\gamma}_{ssc}$ are the cooling rates of synchrotron and inverse Compton emission. q_e is a normalization factor. $\Im_{e,inj}(\gamma, t)$ is the injection spectrum assumed to be a power law with exponential cutoff. The approximate solutions of the differential equation given above are [KTK02]:

$$n_e(\gamma) = \begin{cases} q_e t_{e,esc} \gamma^{-p} & \text{for } \gamma_{min} \leq \gamma < \gamma_{br} \\ q_e t_{e,esc} \gamma_{br} \gamma^{-p-1} & \text{for } \gamma_{br} < \gamma \leq \gamma_{max} \end{cases} \tag{7.3}$$

Thus, the electron population can be phenomenologically described by a broken power-law with minimal and maximal energies E_{min} and E_{max} corresponding to the gamma factors γ_{min} and γ_{max} [K$^+$03]:

$$\frac{dN}{dE} \;\; \sim \;\; \begin{cases} E^{-p} & \text{for } E_{min} < E < E_{br} \\ E^{-(p+1)} & \text{for } E_{br} < E < E_{max} \end{cases} \tag{7.4}$$

$$\tag{7.5}$$

From the simplest version of Fermi acceleration a power law index of $p = 2$ is expected [Lon94]. Relativistic jets can yield $p = 2.23$ [Kra03], but here also higher values like $p = 3$ are possible within certain acceleration models [Lon94]. It is furthermore assumed that the magnetic field inside the spherical volume is of uniform strength and randomly oriented. The break in the spectral index of the e^--distribution at $E = E_{br}$ is due to radiative losses. Figure 7.10 shows schematically the energy spectra of the electron population and the produced synchrotron and inverse Compton radiation.
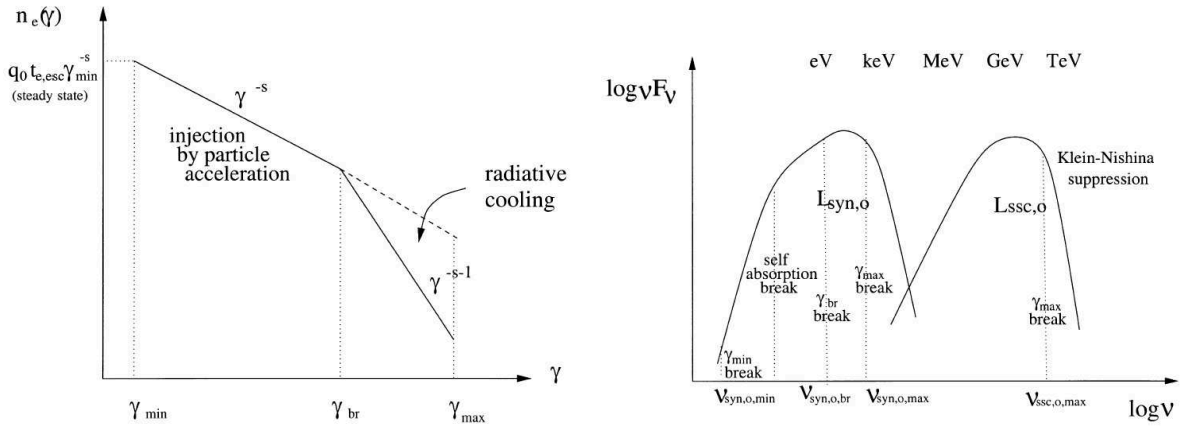
Figure 7.10: Left: The parent electron distribution. The power-law index introduced as p in the text is here denoted as s. $q_0 t_{e,esc} \gamma_{min}^{-s}$ is an approximate description of the electron number density $n_e(\gamma)$ at $\gamma = \gamma_{min}$ (see text). **Right:** Synchrotron and inverse Compton radiation emitted by the parent e^--population within the frame of a simple one-zone Synchrotron Self Compton (SSC) model. $L_{syn,o}$ and $L_{ssc,o}$ are the observed bolometric luminosities of the synchrotron and inverse Compton components. See text for further explanations. Graphics taken from [KTK02].

The break at energy E_{br} (or, correspondingly, at the Lorentz-factor γ_{br}) is determined by the electron radiative cooling time $\tau_{cool}(\gamma)$ being equal to the adiabatic (particle) loss time τ_{ad} [IT96]. Radiative cooling then typically leads to a break by one power in the index of the e^- power law distribution. The break occurs, since τ_{cool} depends on the particle's energy while τ_{ad} is an energy independent particle loss time.

The injected e^--population emits synchrotron and inverse Compton radiation, which (to some degree) mirror the features of the e^--spectrum mentioned above. Therefore, γ_{min}, γ_{max} and γ_{br} are observable in the synchrotron spectrum. Yet, here another feature - the so-called synchrotron self absorption - leads to a further break (see figure 7.10), when the synchrotron emission volume's related luminosity exceeds a threshold [Lon92]. The inverse Compton emission break in the Klein-Nishina limit occurs when the energy transferred to the Compton scattered γ becomes equal to the electron energy E_e and $E_e < E_{max}$. The following parameters, as partly described above, enter into the one-zone SSC model:

- *Redshift z of the source*
  z determines the extinction of the IC emission due to e^+e^- -pair production with EBL target photons, but also (as distance) the measured (apparent) source luminosity.

- *Doppler factor δ*
  The γ-radiation is Doppler-shifted by the factor

$$\delta = \frac{1}{\gamma(1 - \beta cos\theta)} \tag{7.6}$$

Here, $\beta = v/c$ denotes the bulk velocity of the plasma (normalized to the speed of light), θ is the observer's viewing angle and γ is the Lorentz factor.

- *Magnetic field strength B*
  The magnetic field **B** of uniform strength and random orientation inside the spherical emission volume determines the synchrotron luminosity.

- *Radius R of emission volume*

- *Energy density ρ_e of the electron population*

- *E_{min}, E_{max} and E_{br}*
  E_{br} as determined by $\tau_{cool} = \tau_{ad}$ (see above) is fully defined by the magnetic field energy density u_B, the soft photon energy density u_{soft}, the emission region size R and the source region's velocity β (see [IT96]):

$$\tau_{cool} \;=\; \frac{3mc}{4(u_B + u_{soft})\sigma_T\gamma} \tag{7.7}$$

$$\tau_{ad} \;=\; \frac{R}{c} \tag{7.8}$$

In the literature, one can often find the *ratio η between e^--energy density and magnetic field energy density* . The magnetic field energy density can be written as (using a conversion to SI units)

$$u_B = \left(\frac{B^2}{8\pi}\right)_{\text{Gaus}} = \left(\frac{B^2}{2\mu_0}\right)_{\text{SI}} , \; \mu_0 = 4\pi \cdot 10^{-7} \, \text{NA}^{-2} \tag{7.9}$$

$$u_B = \frac{1}{8\pi}\left(\frac{B}{1\,\text{T}}\right)^2 10^8 \frac{\text{erg}}{\text{cm}^3} \tag{7.10}$$

Thus, an evaluation of $\eta = \frac{\rho_e}{u_B}$ for ρ_e given in units of erg/cm^3 is straightforward.

It shall be noted here, that the presented one-zone SSC model is not expected to correctly model the radio to optical fluxes - it underestimates them. The reason is, that radio to optical radiation is likely to be dominated by emission from other regions than those that emit the bulk of the X-rays and γ-rays [IT96]. Especially thermal emission originating from the accretion disk or star light from the host galaxy may contribute.

7.3.1   Constraining the redshift

In order to derive constraints on the redshift of PG1553+113 a 'minimal' EBL density model was chosen (see figure 7.11) by scaling down the EBL 'best fit' model provided by Kneiske [KBMH04], so that the hard limits as derived from galaxy counts are reached. The spectrum, as shown in 7.7, was absorption corrected for different assumed redshifts z. For a given redshift the $\gamma\gamma$-absorption was calculated according to the method described in section 1.5.
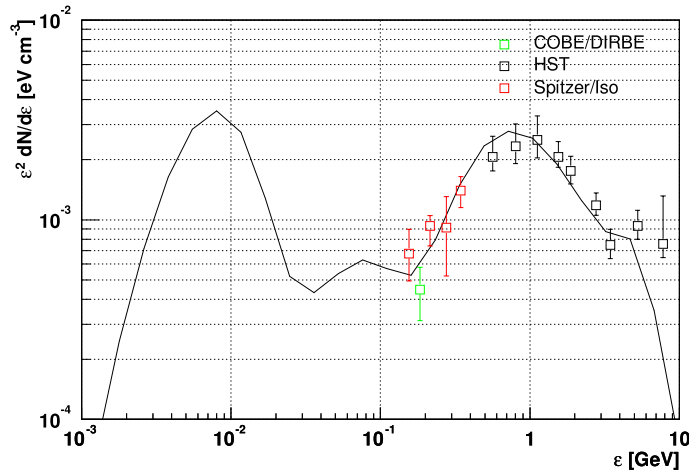
Figure 7.11: The EBL density - model and data. The data are integrated light from galaxy counts (taken from [HD01]) and can therefore be regarded as a hard EBL-limit. The EBL model was taken from [KBMH04] and scaled down to a 'minimal' EBL level by applying a scaling factor of 0.6 $\varepsilon^2 \frac{dn}{d\varepsilon}$ to the original 'best-fit' model, which approximately corresponds to a χ^2-fit to the data. It shall be noted that due to the threshold energy for photon-photon absorption only the second peak in the EBL density is decisive for VHE photon energies E_{VHE} in the range $100\,\text{GeV} < E_{VHE} < 1000\,\text{GeV}$ (compare figure 1.9).
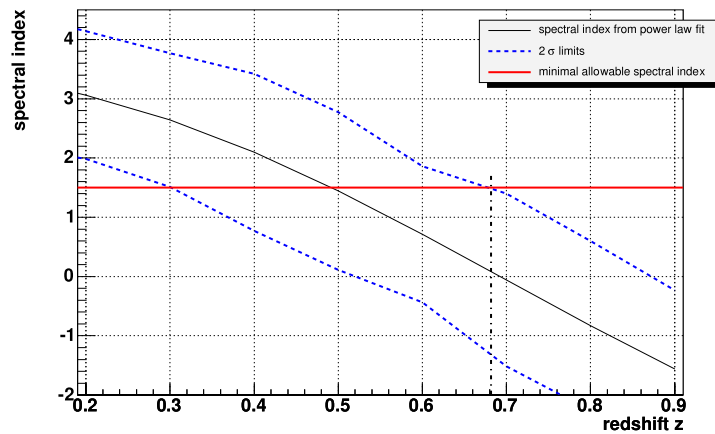
The absorption corrected spectra were fit by means of a power-law function not taking into account the first energy bin since the corresponding flux value showed systematic fluctuations depending on the g/h-separation cuts (see the Crab Nebula analysis as presented in chapter 6). In addition to the statistical error of the flux values F_i of each energy bin i a systematic error was taken into account. Therefore, the flux deviation $\Delta_{F,i} = F_{a,i} - F_{b,i}$ was evaluated with the indices a and b indicating the flux as derived in this work and the flux as presented in [A+06c]. Under the assumption that $\Delta_{F,i}$ roughly corresponds to a 1σ deviation it was taken as the systematic error of the flux value in bin i.

The figures 7.12 and 7.13 show the spectral index as a function of the assumed redshift z. The dashed lines represent the 2σ lower and upper limits. These limits were calculated taking into account the statistical (fit) error of the spectral index. Figure 7.13 shows the result of the same approach but omitting the last two data points of the differential energy flux. Due to their large statistical and systematic errors there is no significant change in comparison with figure 7.12, where all data points have been used. Considering a minimal spectral index of $\alpha = 1.5$ [A+06a] for the intrinsic spectrum, a redshift limit of $z \lesssim 0.68$ can be derived.

This limit seems to violate the lower limit $z > 0.78$ as given in [STF05] (there is consistency for a 3σ upper limit). But one must also take into account the possibilities, that the assumed typical host galaxy's magnitude m_H may not be applicable and that the method given in [STF05] may show unknown systematic errors due to selection effects of the HST survey sample.
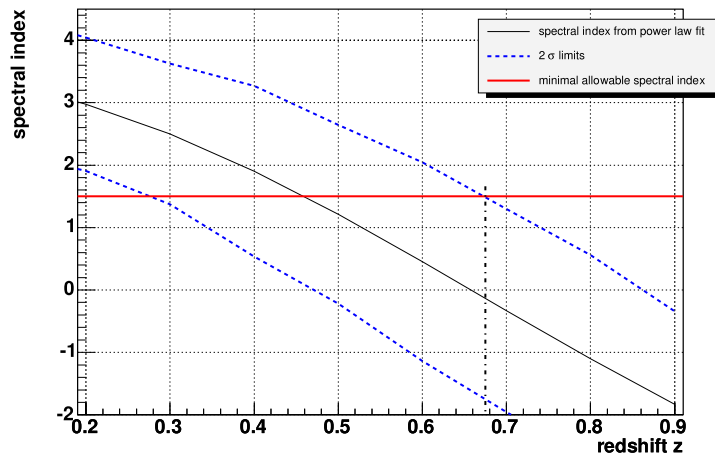
Figure 7.12: The spectral index obtained after a fit to the absorption corrected spectra for different assumed redshifts. The two dashed lines correspond to the 2σ upper and lower limits, the solid black line denotes the spectral index fit value and the solid red line is the minimal physically reasonable spectral index [A+06b].

Figure 7.13: The spectral index obtained after a fit to the absorption corrected spectra for different assumed redshifts. The last two points of the differential energy flux have been excluded in order to judge their influence on the spectral index. See figure 7.12 for further information.

The limit $z < 0.78$ as presented in [A$^+$06c] shows (exactly) a consistency within the 2σ level with [STF05]. In the H.E.S.S. analysis of PG1553+113 presented in [A$^+$06b] a 2σ upper limit of $z < 0.74$ is derived.

7.3.2   The spectral energy distribution

Since up to now only blazars with redshift $z < 0.3$ have been detected in the VHE-γ energy range it seems reasonable to investigate the spectral energy distribution (SED) of PG1553+113 assuming in the beginning $z = 0.3$, which is the lowest possible value according to a simple redshift estimate based merely on the fact that the host galaxy is unresolved in HST surveys (see discussion at the beginning of this chapter). This allows furthermore a comparison with the SED model given in [CG02], where a redshift of $z = 0.36$ was assumed, and also a comparison with [A$^+$06c], where $z = 0.3$ was taken.

For the investigation of the PG1553+113 SED the mentioned one-zone SSC model was fit to optical, X-ray and γ-ray data. The X-ray measurements are from BeppoSAX observations [G$^+$02] and the EGRET upper limit was taken from the first EGRET catalogue [F$^+$94]. The radio and optical data are available through the URL given in [G$^+$02], where it is indicated that the optical data are selected from the GSC-II catalog [CAS06]. The average optical flux as observed by the KVA telescope on La Palma simultaneously to the γ-ray observations shows a comparable though a little higher flux level [A$^+$06c].

The figures 7.14 and 7.15 show the SEDs for the 'best-fit' and the minimum-EBL model, respectively. The red curve depicts the synchrotron SED whereas the blue curve shows the distribution of the Compton-upscattered VHE photons as expected in a SSC scenario (Inverse Compton, or shortly IC SED). Both the intrinsic IC SED (green curve) and the distribution of the VHE photons corrected for extragalactic extinction (photon-photon absorption, blue curve) are shown. The corresponding fit-parameters can be found in table 7.2.

The radio data were not taken into account for the fit since the radio emission is supposed to originate from a much larger volume than that which is used within the one-zone SSC model. However, the optical data can be fit reasonably well within the model. The absence of emission lines, the unresolved host galaxy as well as the high degree of polarization [A$^+$06c] indicate that the optical emission is indeed synchrotron radiation *without* a significant contribution from starlight and thermal accretion disk radiation.

The parameters listed in table 7.2 are in general agreement with [CG02], where radio to X-ray data from observations of PG1553+113 were exploited. In the independent analysis [A$^+$06c] the following parameter values are found (the same SSC simulation code is used): Doppler factor $\delta = 15$, magnetic field $B\,[\mathrm{G}] = 0.33$, radius of the emitting region $R\,[10^{14}\,\mathrm{cm}] = 3$, electron energy density $\rho_e\,[\mathrm{erg/cm^3}] = 0.04$, minimum energy $log_{10}(E_{min}/\mathrm{eV}) = 8$, break energy $log_{10}(E_{br}/\mathrm{eV}) = 10.2$, maximum energy $log_{10}(E_{max}/\mathrm{eV}) = 11$, photon index of the electron distribution for energies below the break energy $p = 2.6$. These values lead to a ratio $\eta = \frac{\rho_e}{u_B} = 9$, somewhat smaller than what is found in this work.

The SED fit for $z = 0.6$ shows a significantly higher χ^2-value of $25 - 34$ for photon-photon absorption by a minimal EBL. A 'best-fit' EBL even results in χ^2-values larger than 40. This finding suggests that the upper redshift limit is even lower than $z < 0.68$.

redshift $z = 0.3$ 'best fit' EBL
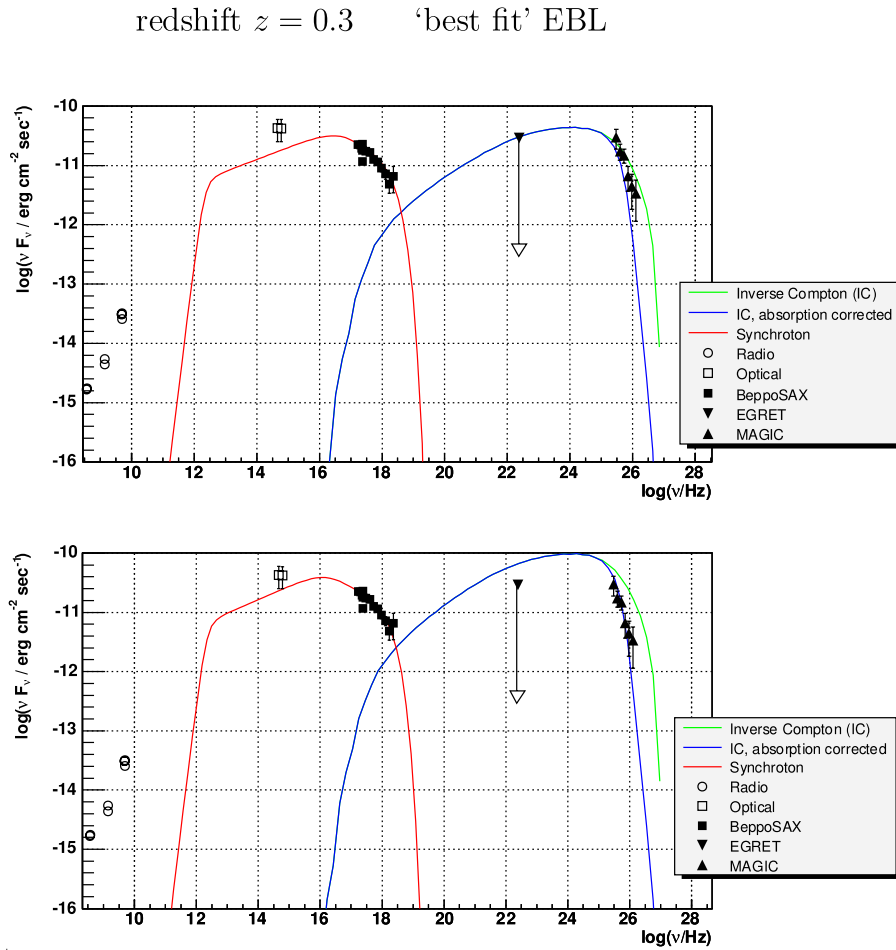
Figure 7.14: Model fit to experimental data. The radio data were not taken into account, since they are expected not to be properly described by the one-zone SSC model (see text). The X-ray data are from BeppoSAX observations [G+02], the EGRET upper limit was taken from the first EGRET catalogue [F+94]. The MAGIC flux estimates are derived from the analysis presented in this chapter. A 'best fit' EBL density model was used. **Top:** SED fit using the EGRET upper limit as constraint ($\chi^2/n_{dof} = 30/12$). **Bottom:** SED fit without consideration of the EGRET upper limit ($\chi^2/n_{dof} = 14.6/12$).

redshift $z = 0.3$ minimal EBL
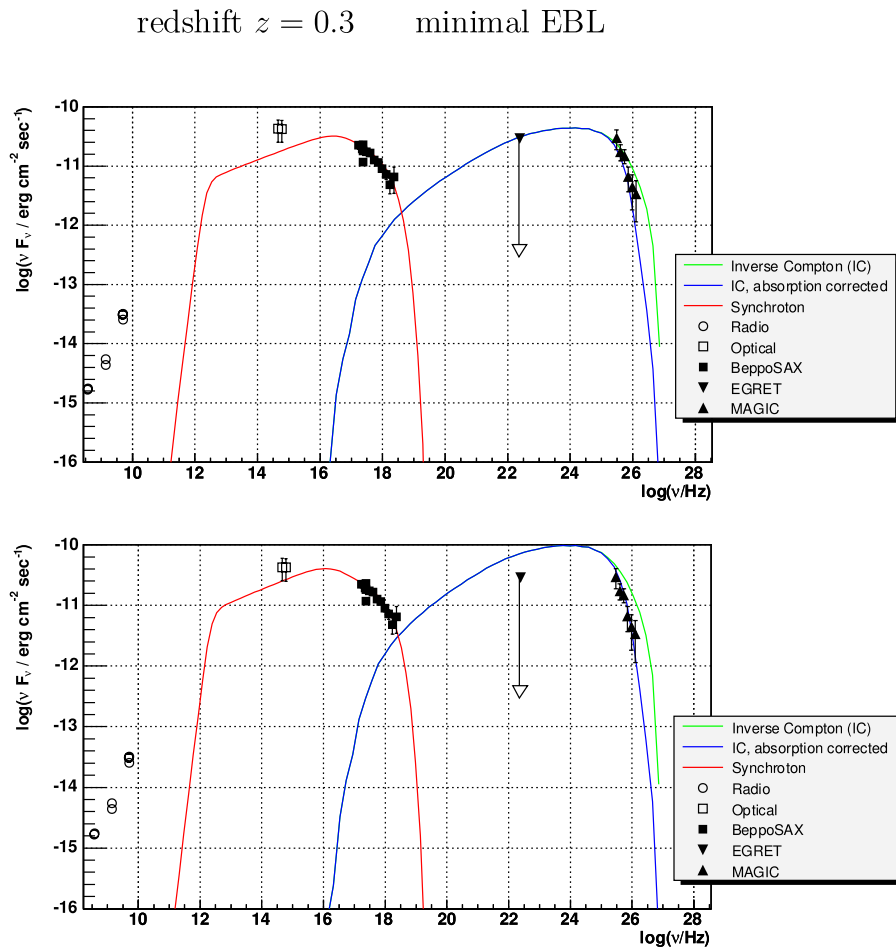
Figure 7.15: Model fit to experimental data. See figure 7.14 for details. A 'minimal' EBL density model was used. **Top:** SED fit using the EGRET upper limit as constraint ($\chi^2/n_{dof} = 21.5/12$). **Bottom:** SED fit without consideration of the EGRET upper limit ($\chi^2/n_{dof} = 10.6/12$).

	EBL_{min} woUL	EBL_{min} wUL	EBL_{best} woUL	EBL_{best} wUL
Doppler factor δ	18.5 ± 0.1	18.1 ± 0.3	23 ± 2	18.5 ± 0.1
magn. field B [G]	0.211 ± 0.007	0.25 ± 0.05	0.2 ± 0.1	0.25 ± 0.08
radius R [10^{14} cm]	2.13 ± 0.03	2.07 ± 0.06	1.7 ± 0.3	2.07 ± 0.06
ρ_e [10^{-2} erg/cm^3]	9.3 ± 0.4	5.7 ± 0.5	11 ± 3	5.7 ± 0.5
$log_{10}(E_{min}/\text{eV})$	8.09 ± 0.02	8.10 ± 0.2	8.11 ± 0.1	8.10 ± 0.2
$log_{10}(E_{br}/\text{eV})$	10.20 ± 0.02	10.36 ± 0.04	10.2 ± 0.2	10.36 ± 0.03
$log_{10}(E_{max}/\text{eV})$	11.18 ± 0.01	11.18 ± 0.01	11.2 ± 0.1	11.18 ± 0.01
photon index p	2.58 ± 0.01	2.59 ± 0.01	2.5 ± 0.1	2.59 ± 0.01
u_B[10^{-2} erg/cm^3]	0.18 ± 0.01	0.25 ± 0.01	0.10 ± 0.01	0.25 ± 0.01
$\eta = \frac{\rho_e}{u_B}$	53 ± 2	23 ± 1	112 ± 27	23 ± 1
χ^2/n_{dof}	$10.6/12$	$21.5/12$	$14.6/12$	$30/12$

Table 7.2: Fit results for non-simultaneous PG1553+113 multiwavelength observations. EBL_{min} denotes the minimum EBL level as discussed in this section and EBL_{best} denotes the 'best-fit' EBL model [KBMH04]. The abbreviations 'woUL' and 'wUL' indicate a fit without and with consideration of the EGRET upper limit, respectively. A redshift $z = 0.3$ of the source was assumed.

Furthermore, as already mentioned elsewhere (see [A$^+$06a] and [KKT05]), a minimal EBL level seems to be in better agreement with the observed SEDs of high-redshift sources. The same is observed here, when the EGRET upper limit is taken into account, already for a assumed redshift of $z = 0.3$.

In [KKT05] the blazar H1426+428 is investigated ($z = 0.129$). For best-matching low-level EBL models the authors state that their SSC model parameters are in agreement with what is known from the analysis of Mkn 421 and Mkn 501. However, they find high values of $\eta = 190$, which is much larger than the values of about 10 for Mkn 421 and Mkn 501. In the analysis of PG1553+113 as presented in this work, this trend is not so dramatic, but values $\eta \gtrsim 23$ are already significantly higher than those for the two mentioned 'Markarian blazars'. An energy density of the non-thermal electrons which is dominating over the magnetic field energy density seems to be rather typical (see e.g. [K$^+$03, Kra03, IT96]).

Chapter 8

Conclusion and Outlook

In this thesis new data analysis methods especially suited for the investigation of low energy γ-ray events were implemented, successfully tested by means of MC studies and applied to observational data of the Crab Nebula and of the BL Lac object PG1553+113. These methods reach from image cleaning techniques and the utilization of new image parameters to sophisticated g/h-separation and energy estimation approaches. For the first time in γ-ray astronomy the advantages of classification and regression trees were exploited in order to improve existing 'classical' methods.

In the analysis of observational data from the BL Lac PG1553+113 significant excess for both years 2005 and 2006 could be obtained, the combined data showing a signal in excess of 8σ. In the further analysis a spectrum could be derived for the merged data sets of 2005 and 2006. This spectrum could be utilized to constrain the redshift z of PG1553+113 with the result $z \lesssim 0.68$. The broad-band SED could be fit reasonably well using a simple SSC model. Fits for different assumed redshifts and EBL density levels lead to the following findings:

A minimal EBL density seems to be in better agreement with the observed SED of PG1553+113. The upper redshift limit for PG1553+113 seems even lower than $z < 0.68$.

The PG1553+113 analysis presented in this thesis shows results, which are in reasonable agreement with those from an independent analysis [A+06c], which uses a different image cleaning but the same g/h-separation and energy estimation methods. However, following sources of systematic errors must be considered:

The upper limit estimate for the redshift and the IC SED as shown in the multi-wavelength analysis depend on the treatment of the photon-photon absorption. Here, uncertainties in the cosmological model play some role. In this work the so-called concordance model with parameters $\Omega_\lambda = 0.7$ and $\Omega_M = 0.3$ was used, which is the established model of the last few years.

Moreover, an intrinsic spectrum with photon index Γ smaller than 1.5 may be possible. Under certain assumptions $\Gamma = 1.25$ can be found, though the corresponding scenario of ineffective radiative cooling together with the IC emission being wholly within the Thomson limit seems unlikely [A+06a].

To a larger extent, uncertainties concerning the multi-wavelength SED fit arise since the considered data were not taken simultaneously, but rather even at different years.

Especially the EGRET upper limit may not be a representative threshold for the real flux level in this energy range. Apart from that, a simple one-zone SSC scenario was taken as model which may not provide an adequate description of the γ-ray and VHE-γ emission of PG1553+113. There may be a significant contribution of external seed photons involved in the IC scattering. Intrinsic absorption was not taken into account.

As outlook one can say that *simultaneous* multi-wavelength observations especially in the former EGRET range can help to constrain the SSC model parameters and therefore to provide estimates of the redshift and the EBL density level. Furthermore, a detection of PG1553+113 at TeV energies would be very useful, since the differential energy flux then should show a more characteristic absorption 'signature' (compare figure 1.12).

Appendix A

Agreement between Monte Carlo Simulations and observational data - supplementary plots

See section 6.2 for a better understanding of the following plots.

$dist$ vs $\log_{10}(size)$
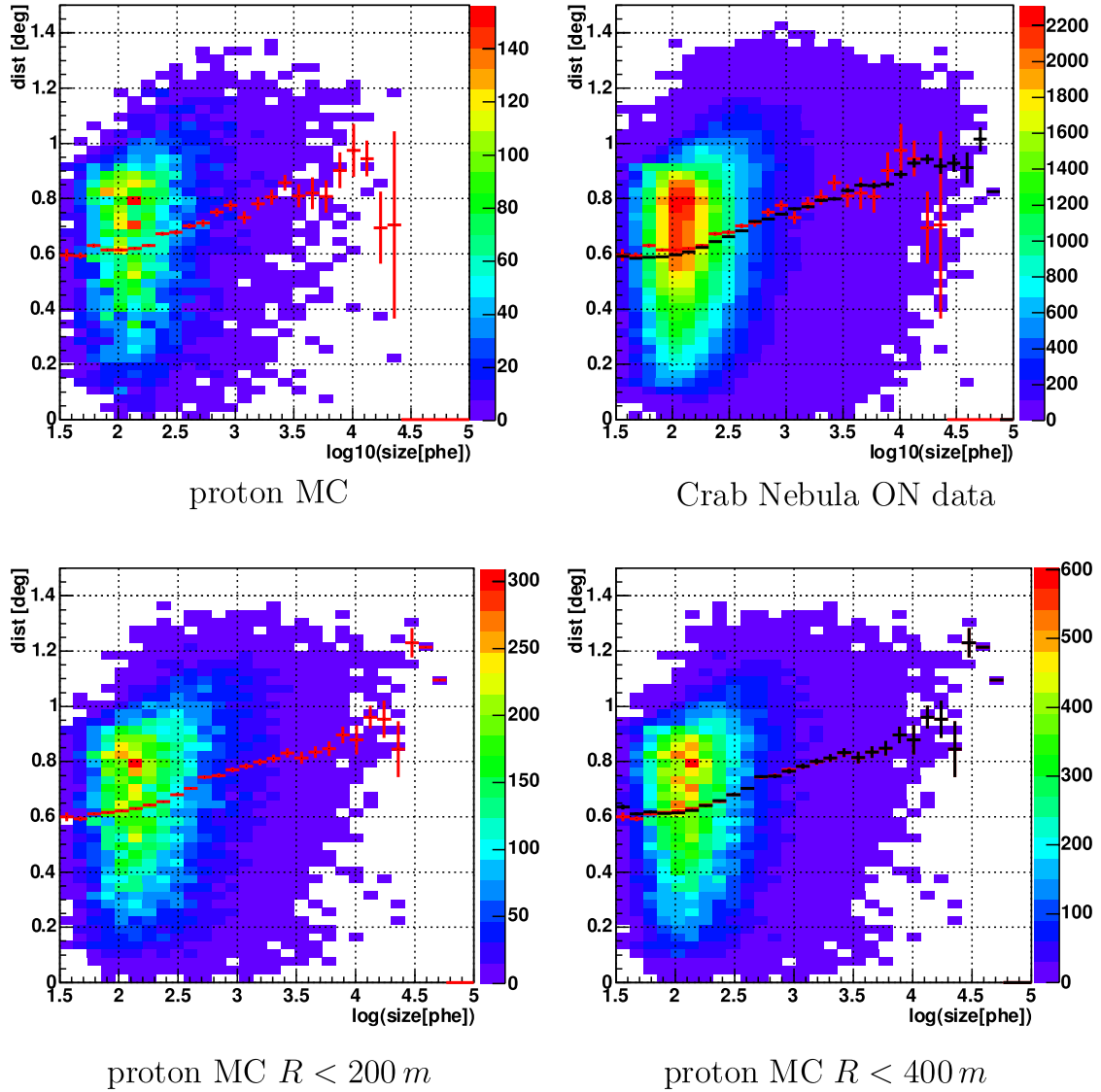
Figure A.1: **Top:** *dist vs* $\log_{10}(size)$ plot for the *p*-MC (histogram on the left and red profile plot) and Crab Nebula ON data (histogram on the right and black profile plot). **Bottom:** *dist vs* $\log_{10}(size)$ plot of the *p*-MC for impact parameters $R < 200$ m (histogram on the left and red profile plot) and for impact parameters $R < 400$ m (histogram on the right and black profile plot). Since no significant changes are introduced by a limited impact parameter range, the good agreement documented in the plots on the top can be taken as reliable.

$length$ vs $\log_{10}(size)$

proton MC

Crab Nebula ON data

proton MC $R < 200\,m$

proton MC $R < 400\,m$
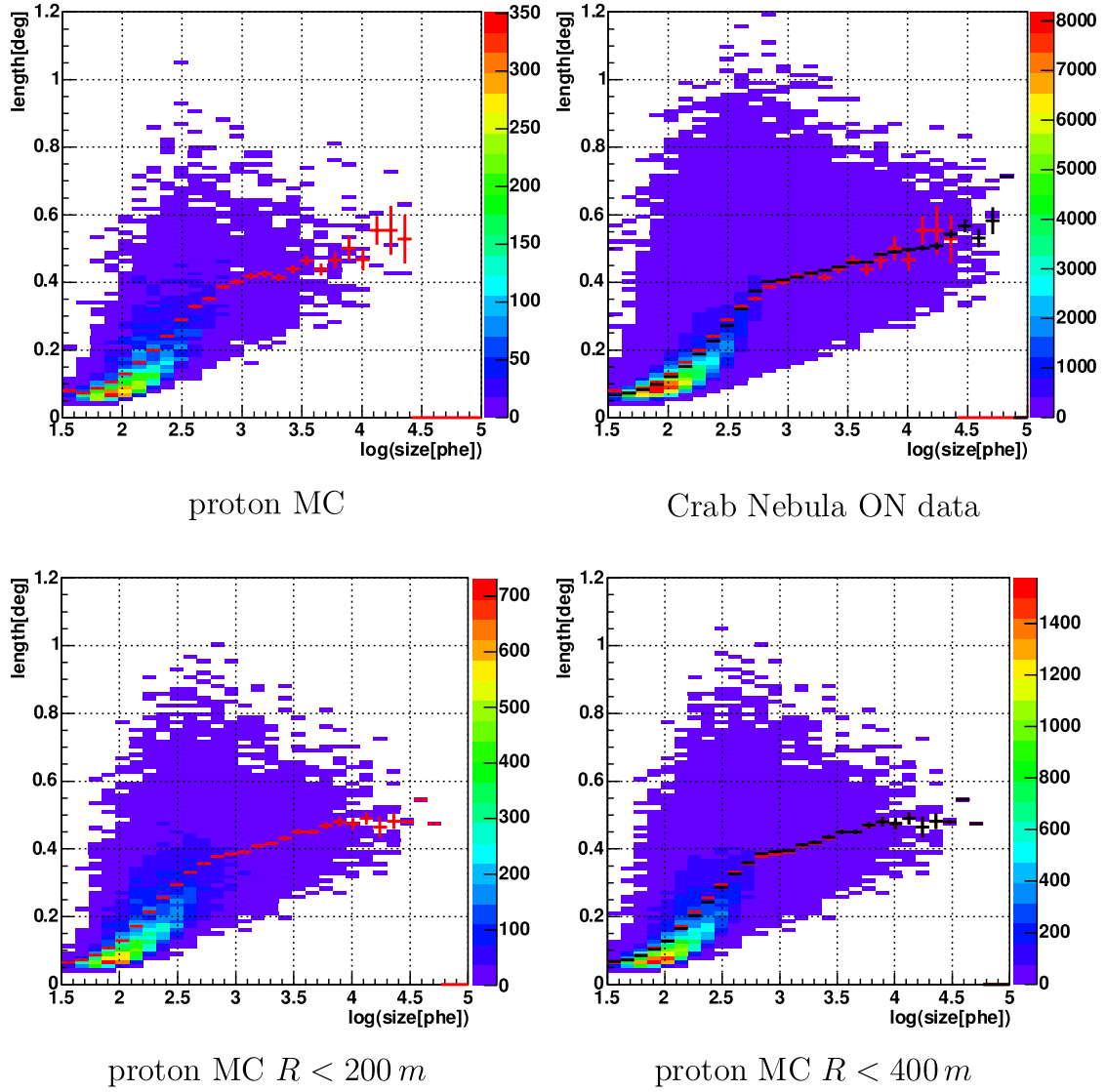
Figure A.2: $length$ versus $\log_{10}(size)$ histograms. See figure A.1 for further explanations.

relative island size   *cosine of charge asymmetry angle*
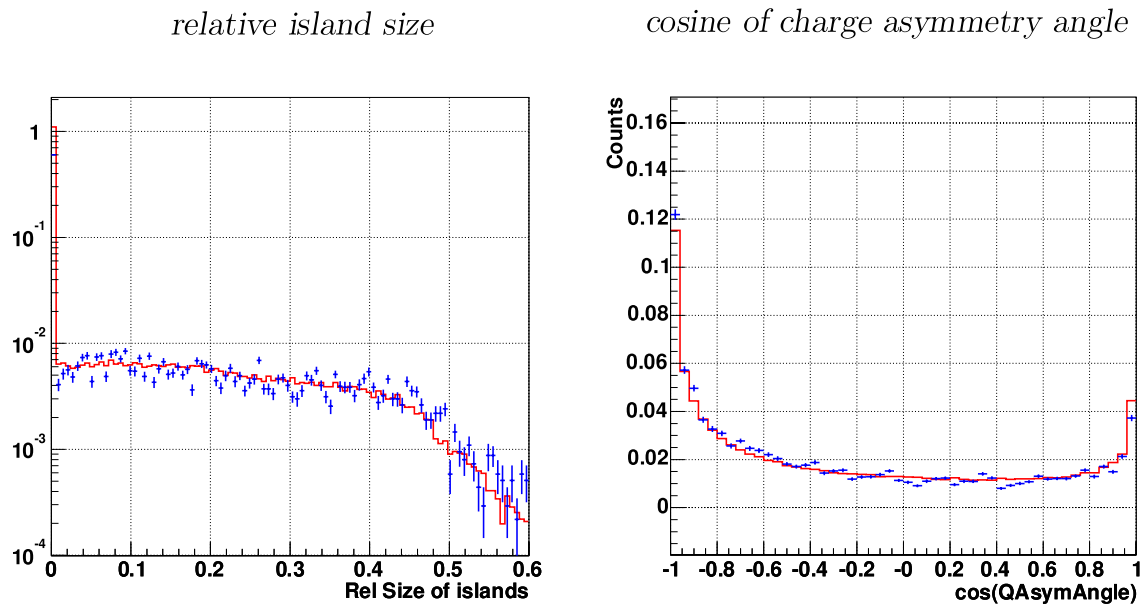
Figure A.3:   Comparison between Crab Nebula ON data (blue data points) and MC protons (red histogram) for the following image parameter distributions. **Left:** Histogram of the *relative island size* for the whole *size* range. **Right:** Histogram of the cosine of the charge asymmetry angle for the whole *size* range. The asymmetric shape probably has its origin in optical aberrations, which increase with increasing 'camera center' to 'image center' distance and produce a radial image tail.

$dens$ vs $\log(size)$

proton MC

Crab Nebula ON data

proton MC $R < 200\,m$

proton MC $R < 400\,m$
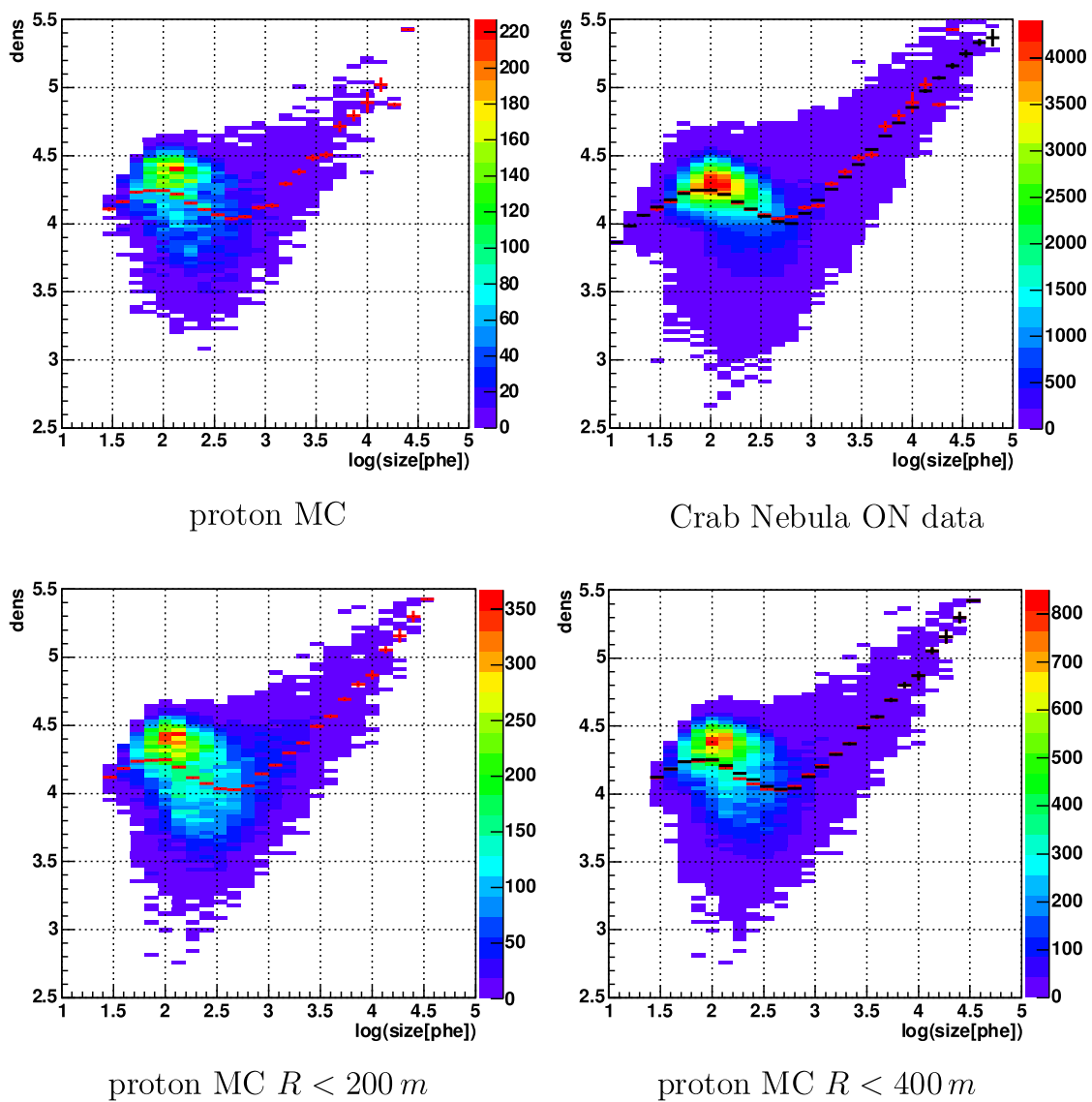
Figure A.4: $dens$ versus $\log_{10}(size)$ histograms. See figure A.1 for further explanations.

Figure A.5:   *relative island size* versus $\log_{10}(size)$ histograms.   See figure A.1 for further explanations.

$$\cos(\measuredangle(\vec{a}_q, \vec{r}_q)) \text{ vs } \log_{10}(size)$$

proton MC

Crab Nebula ON data

proton MC $R < 200\,m$

proton MC $R < 400\,m$
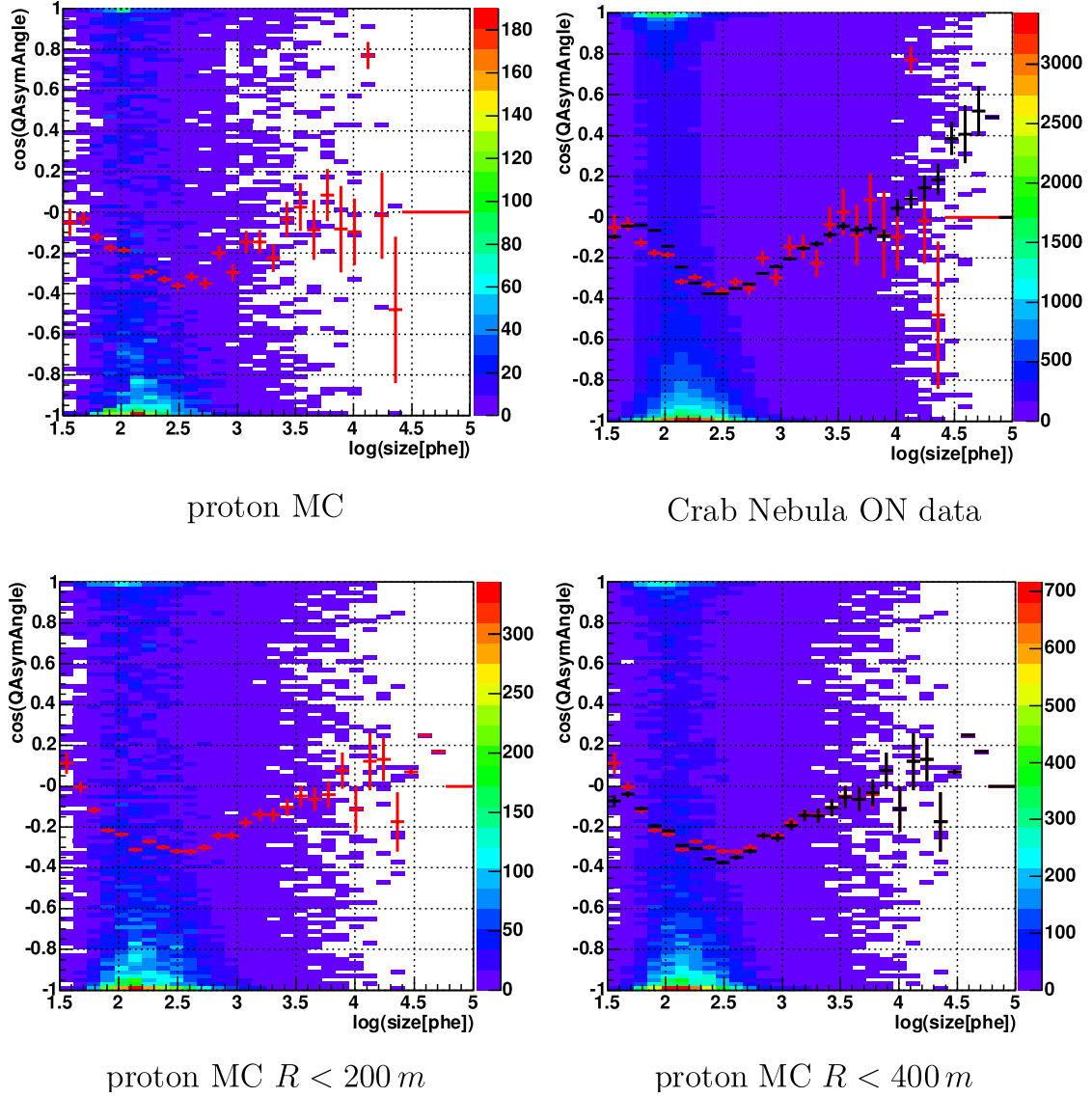
Figure A.6: $\cos(\measuredangle(\boldsymbol{a}_q, \boldsymbol{r}_q))$ versus $\log_{10}(size)$ histograms. See figure A.1 for further explanations concerning the plot set and 4 for details about the asymmetry parameter.
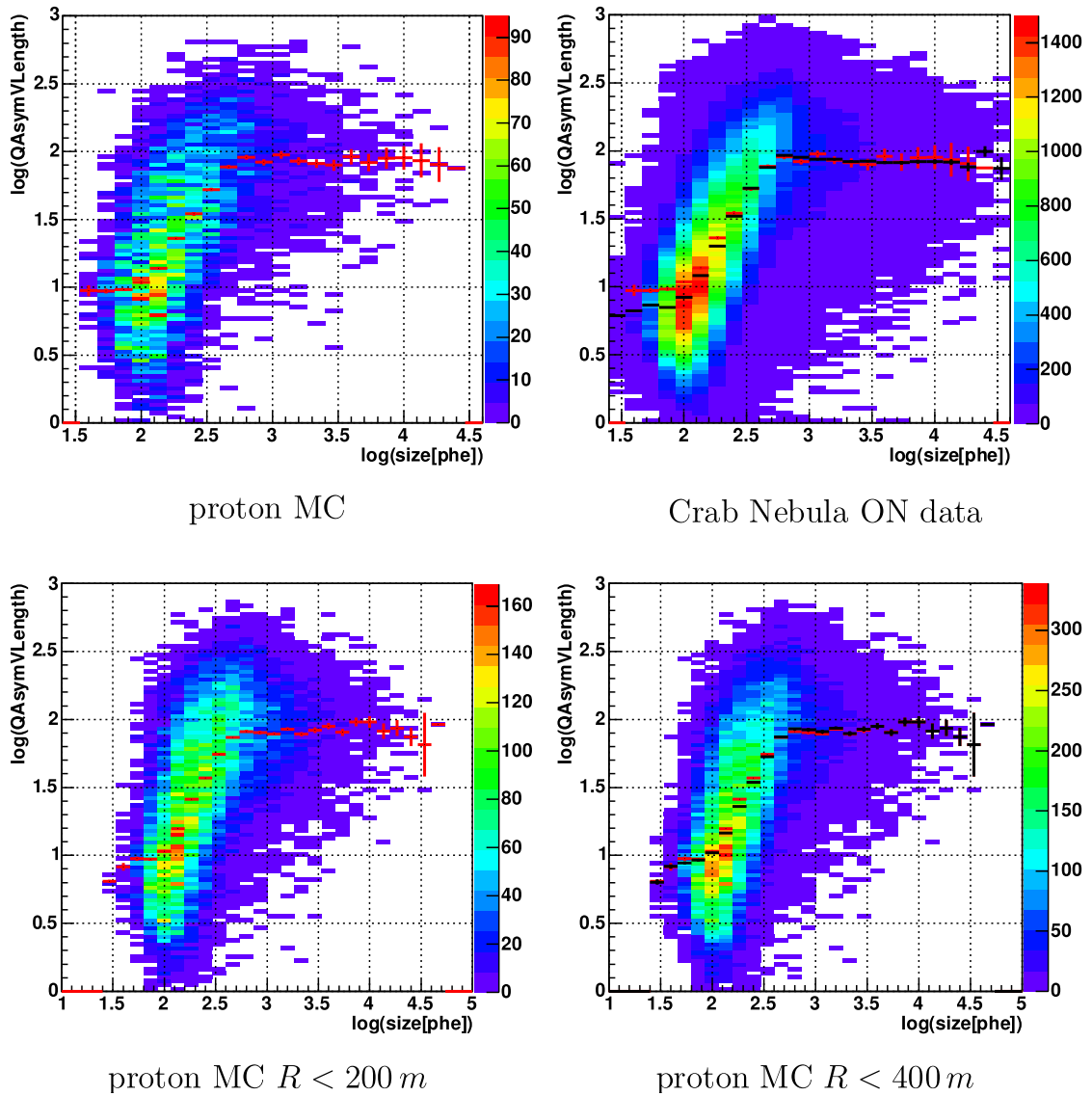
Figure A.7: $length_q$ versus $\log_{10}(size)$ histograms. See figure A.1 for further explanations concerning the plot set and 4 for details about the asymmetry parameter.
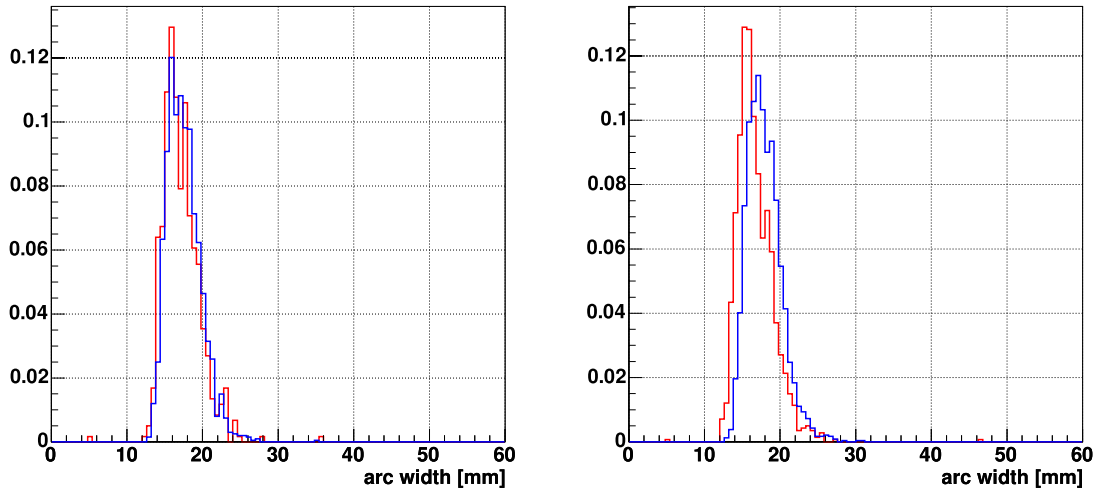
Figure A.8: Left: Muon *arc width* distributions for data taken on PG1553 in 2005 (*red histogram*) and for a MC muon simulation (*blue histogram*). The best MC settings were PSF=7 mm and reflectivity 68 %. **Right:** Significant mismatch for a muon MC with PSF=10 mm (*blue histogram*). See caption of figure 6.3 for further explanations.
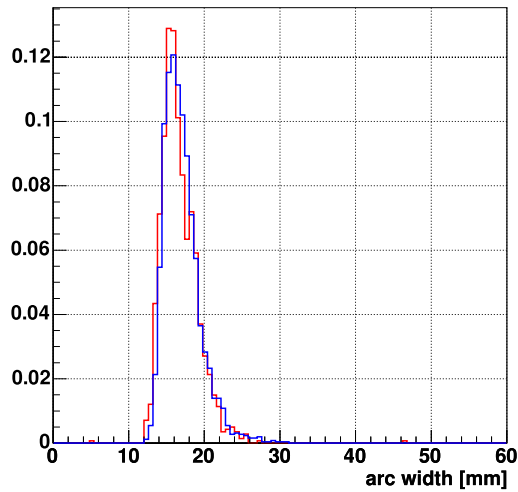
Figure A.9: Muon *arc width* distributions for data taken on PG1553+113 in 2006 (**red histogram**) and MC muon simulation (**blue histogram**). The best MC settings were PSF=10 mm and reflectivity 63 %. See caption of figure 6.3 for further explanations.
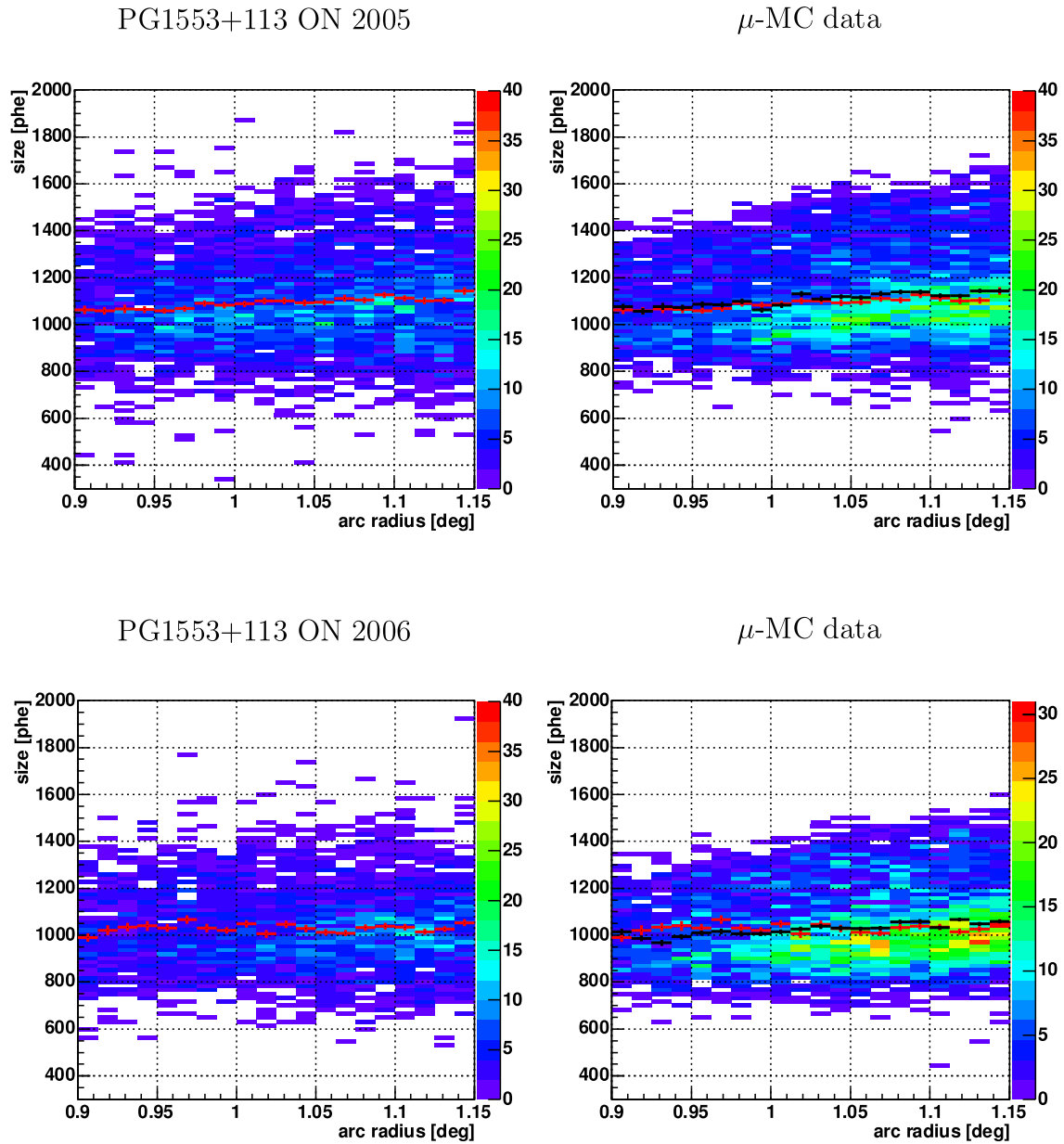
Figure A.10: *size* vs *arc radius* distribution for PG1553+113 ON data and μ-MC. See caption of figure 6.3 for further explanations.

Bibliography

[A+97]   Aharonian, F.A.; et al.: The potential of ground based arrays of imaging atmospheric Cherenkov telescopes. I. Determination of shower parameters. In: *Astroparticle Physics*, volume 6:pp. 343–368, April 1997.

[A+00]   Aharonian, F.A.; et al.: The energy spectrum of TeV gamma rays from the Crab Nebula as measured by the HEGRA system of imaging air Cherenkov telescopes. In: *The Astrophysical Journal*, volume 539:pp. 317–324, August 2000.

[A+04]   Aharonian, F.A.; et al.: The Crab Nebula and pulsar between 500 GeV and 80 TeV: Observations with the HEGRA stereoscopic air Cherenkov telescopes. In: *The Astrophysical Journal*, volume 614:pp. 897–913, October 2004.

[A+05]   Aharonian, F.A.; et al.: H.E.S.S. observations of PKS 2155-304. In: *Astronomy & Astrophysics*, volume 430:pp. 865–875, 2005.

[A+06a]   Aharonian, F.A.; et al.: A low level of extragalactic background light as revealed by γ-rays from blazars. In: *Nature*, volume 440:pp. 1018–1021, 2006.

[A+06b]   Aharonian, F.A.; et al.: Evidence for VHE γ-ray Emission from the distant BL Lac PG1553+113, 2006. astro-ph/0601545.

[A+06c]   Albert, J.; et al.: Detection of VHE radiation from the BL Lac PG1553+113 with the MAGIC Telescope, 2006. astro-ph/0606161, submitted to ApJ Letters.

[AG84]   Allkofer, O.C.; Grieder, P.K.F.: *Cosmic Rays on Earth*. Fachinformationszentrum Energie, Physik, Mathematik GmbH, 1984.

[B+97]   Barrio, J.A.; et al.: The Magic Telescope, Design study for the construction of a 17m Cerenkov telescope for Gamma-Astronomy above 10 GeV. internal report MPI-PhE/98-5, Friedrich-Wilhelms-Universität Bonn, January 1997.

[B+01]   Barate, R.; et al.: Search for supersymmetric particles in e^+e^- collisions at $\sqrt{s}$ up to 202 GeV and mass limit for the lightest neutralino. In: *Physics Letters B*, volume 499:pp. 67–84, 2001. ALEPH Collaboration.

[B⁺04]   Baixeras, C.; et al.: Commissioning and first tests of the MAGIC Telescope.
            In: *Nucl. Inst. Meth. A*, volume 518:p. 188, 2004.

[B⁺05]   Bartko, H.; et al.: Comparison of Signal Reconstruction Algorithms for
            the MAGIC Telescope. internal report TDAS-05-03, MAGIC Collaboration,
            2005.

[BCG⁺04]   Bock, R.K.; Chilingarian, A.; Gaug, M.; Hakl, F.; Hengstebeck, T.; M. Ji-
            rina, J. Klaschka; Kotrc, E.; Savicky, P.; Towers, S.; Vaiciulis, A.; Wittek,
            W.: Methods for multidimensional event classification: a case study using
            images from a Cherenkov gamma-ray telescope. In: *Nucl. Instr. and Meth.
            A*, volume 516:pp. 511–528, 2004.

[BFOS83]   Breimann, L.; Friedmann, J.H.; Olshen, R.A.; Stone, C.J.: *Classification
            and Regression Trees*. Wadsworth, 1983.

[BGG05]   Berezinsky, V.; Gazizov, A.Z.; Grigorieva, S.I.: Dip in UHECR spectrum as
            signature of proton interaction with CMB. In: *Physics Letters B*, volume
            612:pp. 147–153, 2005.

[BL01]   Buja, A.; Lee, Y.S.: Data Mining Criteria for Tree-Based Regression and
            Classification. In: *Proceedings of the seventh ACM SIGKDD international
            conference on knowledge discovery and data mining, San Francisco, Califor-
            nia*, pp. 27–36. 2001.

[Bla04]   Blanch, O.: *Study of the Gamma Ray Horizon with MAGIC as a new method
            to perform cosmological measurements*. Ph.D. thesis, IFAE Barcelona, 2004.

[BM03]   Bastieri, D.; MAGIC Collaboration: The Reflecting Surface of the MAGIC
            Telescope. In: *Proceedings of the 28th International Cosmic Ray Conference,
            Tsukuba, Japan*, pp. 2919–2022. 2003.

[BM04]   Blanch, O.; Martinez, M.: Exploring the Gamma ray Horizon with the next
            generation of Gamma Ray Telescopes. Part 2: Extracting cosmological pa-
            rameters from the observation of γ-ray sources, 2004. astro-ph/0406061.

[BR06]   Brun, R.; Rademakers, F.: ROOT - An Object-Oriented Data Analysis
            framework, 2006. URL `http://root.cern.ch/`.

[Bra92]   Brandt, S.: *Datenanalyse. Mit statistischen Methoden und Computerpro-
            grammen*. BI Wissenschaftsverlag, 1992.

[Bre99]   Breiman, L.: Combining predictors. In: Sharkey, Amanda J.C., editor,
            Combining Artificial Neural Nets. Springer-Verlag, 1999.

[Bre01a]   Breiman, L.: FORTRAN program Random Rorests Version 3.1, 2001. URL
            `ftp://ftp.stat.berkeley.edu/pub/users/breiman`.

[Bre01b]   Breiman, L.: RANDOM FORESTS. Technical report, Statistics Department, University of California, Berkeley, California, 2001. URL `ftp://ftp.stat.berkeley.edu/pub/users/breiman`.

[Bre02]   Breiman, L.: *Manual On Setting Up, Using, And Understanding Random Forests V3.1*, 2002. URL `ftp://ftp.stat.berkeley.edu/pub/users/breiman`.

[Bre06]   Bretz, T.: The MAGIC Analysis and Reconstruction Software (Mars), 2006. URL `http://magic.astro.uni-wuerzburg.de/mars/`.

[BWM03]   Bretz, T.; Wagner, R.; MAGIC Collaboration: The MAGIC Analysis And Reconstruction Software. In: *28th International Cosmic Ray Conference, Tsukuba, Japan*, pp. 2947–2950. 2003.

[CAS06]   The Guide Star Catalog II (GSC-II). The Catalogs and Surveys Group (CASG), 2006. URL `http://www-gsss.stsci.edu/Catalogs/GSC/GSC2/GSC2.htm`.

[CFK⁺02]   Carangelo, N.; Falomo, R.; Kotilainen, J.; Treves, A.; Ulrich, M.-H.: Spectroscopy of BL Lac Objects: new redshifts and mis-identified sources, 2002. astro-ph/0210016.

[CG02]   Costamante, L.; Ghisellini, G.: TeV candidate BL Lac objects. In: *Astronomy & Astrophysics*, volume 384:pp. 56–71, 2002.

[Cor04]   Cortina, J.: Status and first results of the MAGIC Telescope. In: *Astrophysics and Space Science*, volume 297:pp. 245–255, 2004.

[DSG05]   Donato, D.; Sambruna, R.M.; Gliozzi, M.: Six years of BeppoSAX observations of blazars: a spectral catalog. In: *Astronomy & Astrophysics*, volume 433:pp. 1163–1169, 2005.

[EDJ⁺71]   Eadie, W.T.; Drijard, D.; James, F.E.; Roos, M.; Sadoulet, B.: *Statistical Methods in Experimental Physics*. North Holland Publishing Company, 1971.

[Ern06]   Ernenwein, J.P.: NeuNet package for ROOT, 2006. URL `http://e.home.cern.ch/e/ernen/www/NN`.

[Eva55]   Evans, R.D.: *The Atomic Nucleus*. McGraw-Hill, 1955.

[F⁺94]   Fichtel, C.E.; et al.: The first energetic gamma-ray experiment telescope (EGRET) source catalog. In: *The Astrophysical Journal supplement series*, volume 94:pp. 551–581, October 1994.

[Feg97]   Fegan, D.J.: γ/hadron separation at TeV energies. In: *J. Phys. G: Nucl. Part. Phys.*, volume 23:pp. 1013–1060, 1997.

[FT90]   Falomo, R.; Treves, A.: PG 1553 + 11 - A bright optically selected BL
            Lacertae object. In: *Publications of the Astronomical Society of the Pacific
            (PASP)*, volume 102:pp. 1120–1125, October 1990.

[G⁺02]   Giommi, P.; et al.: A Catalog of 157 X-ray Spectra and 84 Spectral Energy
            Distributions of Blazars Observed with BeppoSAX, 2002. astro-ph/0209596,
            data available at given address, URL `http://www.asdc.asi.it/blazars/`.

[Gai90]   Gaisser, T.K.: *Cosmic Rays and Particle Physics.* Cambridge University
            Press, 1990.

[Gal02]   Galant, Y.A.: Particle Acceleration at Relativistic Shocks, January 2002.
            astro-ph/0201243.

[GAM95]   Giommi, P.; Ansari, S.G.; Micol, A.: Radio to X-ray energy distribution of
            BL Lacertae objects. In: *Astron. Astrophys. Suppl.*, volume 109:pp. 267–291,
            1995.

[GH05]   Gillessen, S.; Harney, H.L.: Significance in gamma-ray astronomy - the Li
            & Ma problem in Baesian statistics. In: *Astronomy & Astrophysics*, volume
            430:pp. 355–362, 2005.

[GM05]   Goebel, F.; MAGIC Collaboration: Absolut energy scale calibration of the
            MAGIC telescope using muon images. In: *29th International Cosmic Ray
            Conference, Pune, India*, pp. 101–106. 2005.

[Gre60]   Greisen, K.: Cosmic Ray Showers. In: *Annual Review of Nuclear and Particle
            Sciences*, volume 10:pp. 63–108, 1960.

[Gru93]   Grupen, C.: *Teilchendetektoren.* BI Wissenschaftsverlag, 1993.

[Gru96]   Grupen, C.: *Particle Detectors.* Cambridge University Press, 1996.

[Gru00]   Grupen, C.: *Astroteilchenphysik.* Vieweg, 2000.

[Gru05]   Grupen, C.: *Astroparticle Physics.* Springer, 2005.

[GS67]   Gould, R.J.; Schreder, G.P.: Pair Production in Photon-Photon Collisions.
            In: *Physical review*, volume 155(5), 1967.

[GS02]   Gaug, M.; Schweizer, T.: Calibration analysis for the MAGIC telescope. in-
            ternal note, Institut de Fisica d Altes Energies, university Barcelona, MAGIC
            Collaboration, 2002.

[GS06]   Gaensler, B.M.; Slane, P.O.: The evolution and structure of pulsar wind
            nebulae, 2006. astro-ph/0601081.

[GSL86]   Green, R.F.; Schmidt, M.; Liebert, J.: The Palomar-Green catalog of
            ultraviolet-excess stellar objects. In: *Astrophysical Journal Supplement Se-
            ries (ApJS)*, volume 61:pp. 305–352, June 1986.

[H+98]   Hillas, A.M.; et al.: The spectrum of TeV gamma rays from the Crab Nebula. In: *The Astrophysical Journal*, volume 503:pp. 744–759, August 1998.

[H+99]   Hartman, R.C.; et al.: The third EGRET catalog of high-energy gamma-ray sources. In: *The Astrophysical Journal Supplement Series*, volume 123:pp. 79–202, July 1999.

[HD01]   Hauser, M.G.; Dwek, E.: The Cosmic Infrared Background: measurements and implications. In: *Ann. Rev. Astron. Astrophys.*, volume 39:pp. 249–307, 2001.

[HEA06]   Gamma-Ray Sky Maps. High Energy Astrophysics Science Archive Research Center (HEASARC), 2006. URL `http://heasarc.gsfc.nasa.gov/docs/cgro/images/epo/gallery/skymaps`.

[Hil85]   Hillas, A.M.: In: *Proceedings of the 19th International Cosmic Ray Conference, La Jolla, USA*, volume 3, p. 445. 1985.

[Hil96]   Hillas, A.M.: Differences between gamma-ray and hadronic showers. In: *Space Science Reviews*, volume 75:pp. 17–30, 1996.

[HK02]   Heck, D.; Knapp, J.: *EAS simulation with CORSIKA: A user's manual*, 2002.

[IT96]   Inoue, S.; Takahara, F.: Electron acceleration and gamma-ray emission from blazars. In: *The Astrophysical Journal*, volume 463:pp. 555–564, June 1996.

[K+99]   Kataoka, J.; et al.: High-energy emission from the TeV blazar Markarian 501 during multiwavelength observations in 1996. In: *The Astrophysical Journal*, volume 514:pp. 138–147, March 1999.

[K+03]   Krawczynski, H.; et al.: Multiwavelength Observations of Strong Flares From the TeV-Blazar 1ES 1959+650, 2003. astro-ph/0310158.

[KBMH04]   Kneiske, T.M.; Bretz, T.; Mannheim, K.; Hartmann, D.H.: Implications of Cosmological Gamma-Ray Absorption, II. Modification of gamma-ray spectra. In: *Astronomy & Astrophysics*, volume 413:p. 807, 2004. data available at given address, URL `http://www.physics.adelaide.edu.au/~tkneiske/`.

[KKT05]   Kato, T.; Kusunose, M.; Takahara, F.: Constraining the Emission Properties of TeV Blazar H1426+428 by the Synchrotron-Self-Compton Model, 2005. astro-ph/0510730.

[KNS03]   Kalekin, O.R.; Neshpor, Yu.I.; Stepanyan, A.A.: The Characetristics of Blazars that are Sources of Very-High-Energy Gamma Rays. In: *Astronomy Reports*, volume 47(10):pp. 831–837, October 2003.

[Kon04]   Konopelko, A.: Altitude effect in Ĉerenkov light flashes of low energy gamma-ray-induced atmospheric showers. In: *J. Phys. G: Nucl. Part. Phys.*, volume 30:pp. 1835–1846, 2004.

[Kra01]   Kranich, D.: *Temporal and spectral characteristics of the active galactic nucleus Mkn 501 during a phase of high activity in the TeV range.* Ph.D. thesis, Technische Universität München, 2001.

[Kra03]   Krawczynski, H.: TeV Blazars - Observations and Models, 2003. astro-ph/0309443.

[KTK02]   Kino, M.; Takahara, F.; Kusunose, M.: Energetics of TeV blazars and physical constraints on their emission regions. In: *The Astrophysical Journal*, volume 564:pp. 97–107, January 2002.

[LM83]   Li, T.; Ma, Y.: Analysis Methods for results in gamma-ray astronomy. In: *The Astrophysical Journal*, volume 272:pp. 317–324, September 1983.

[LMKF$^+$99] Laurent-Muehleisen, S.A.; Kollgaard, R.I.; Feigelson, E.D.; Brinkamnn, W.; Siebert, J.: The RGB sample of intermediate BL Lacs, 1999. astro-ph/9905133.

[Lon92]   Longair, M.S.: *High energy astrophysics. Vol. 1 Particles, photons and their detection - 2nd ed.* Cambridge University Press, 1992.

[Lon94]   Longair, M.S.: *High energy astrophysics. Vol. 2 Stars, the Galaxy and the interstellar medium - 2nd ed.* Cambridge University Press, 1994.

[MBFG05]   Masterson, C.; Benbow, W.; Funk, S.; Gillessen, S.: Observations of the Crab Nebula with H.E.S.S. In: *2nd International Symposium on High Energy Gamma-Ray Astronomy, AIP Conference Proceedings*, pp. 617–621. 2005.

[MBH$^+$03] Mase, K.; Bock, R.K.; Hengstebeck, T.; Kalekin, O.; Mirzoyan, R.; Shinozaki, K.; Teshima, M.; MAGIC Collaboration: Isolated Muon Study for the MAGIC Telescope. In: *Proceedings of the 28th International Cosmic Ray Conference, Tsukuba, Japan*, pp. 2955–2958. 2003.

[Mey05]   Meyer, M.: Calibration of the MAGIC telescope using muon rings. internal report TDAS-05-10, MAGIC Collaboration, 2005.

[MG83]   Miller, H.R.; Green, R.F.: Optical/UV Observations of the BL Lac Object PG 1553+11. In: *Bulletin of the American Astronomical Society (BAAS)*, volume 15:p. 957, 1983.

[Mir00]   Mirzoyan, R.: Conversion Factor Calibration for MAGIC Based on the Use of Measured F-Factors of PMTs. internal memo, MAGIC Collaboration, April 2000.

[MKU01]   Meier, D.L.; Koide, S.; Uchida, Y.: Magnetohydrodynamic Production of Relativistic Jets. In: *Science 5*, volume 291(5501):pp. 84–92, January 2001.

[MW86]   Murthy, P.V. Ramana; Wolfendale, A.W.: *Gamma-ray astronomy.* Cambridge University Press, 1986.

[O+01]   Oser, S.; et al.: High-energy gamma-ray observations of the Crab Nebula and pulsar with the Solar Tower Atmospheric Cerenkov Effect Experiment. In: *The Astrophysical Journal*, volume 547:pp. 949–958, 2001.

[Ong98]   Ong, R.A.: Very high-energy gamma-ray astronomy. In: *Physics Reports*, volume 305:pp. 93–202, 1998.

[Ong05]   Ong, R.A.: γ-ray Science and Techniques. In: *Rapporteur talk at the 29th International Cosmic Ray Conference, Pune, India*. 2005.

[P+05]   Perlman, E.S.; et al.: Intrinsic Curvature in the X-Ray Spectra of BL Lacertae Objects. In: *The Astrophysical Journal (ApJ)*, volume 625(2):pp. 727–740, 2005.

[Pan04]   Paneque, D.: *The Magic Telescope: development of new technologies and first observations*. Ph.D. thesis, MPI Munich, 2004.

[Per03]   Perkins, D.: *Particle Astrophysics*. Oxford University Press, 2003.

[Pet97]   Petry, D.: *Beobachtung hochenergetischer γ-Strahlung ($E > 1\,TeV$) aus Richtung der aktiven Galaxien Mkn421 und Mkn 501*. Ph.D. thesis, Technische Universität München, 1997.

[R+04]   Riess, A.G.; et al.: Type Ia supernova discoveries at $z > 1$ from the Hubble Space Telescope: Evidence for past deceleration and constraints on dark energy evolution. In: *The Astrophysical Journal*, volume 607:pp. 665–687, June 2004.

[Sch01]   Schoenfelder, V., editor: *The Universe in Gamma Rays*. Springer-Verlag Berlin Heidelberg, 2001.

[Sch02]   Schweizer, T.: *Analysis of the Large Gamma Ray Flares of Mkn421 as observed with HEGRA CT1 on the island La Palma in 2001*. Ph.D. thesis, IFAE Barcelona, 2002.

[SdJ95]   Stecker, F.W.; de Jager, O.C.: Absorption of High Energy Gamma-Rays by Low Energy Intergalactic Photons, 1995. astro-ph/9501065.

[Sha05]   Shayduk, M.: private communication, 2005.

[SHK+05a]   Shayduk, M.; Hengstebeck, Th.; Kalekin, O.; Pavel, N.A.; Schweizer, Th.: A New Image Cleaning Method for the MAGIC Telescope. In: *29th International Cosmic Ray Conference, Pune, India*, pp. 101–104. 2005.

[SHK+05b]   Shayduk, M.; Hengstebeck, Th.; Kalekin, O.; Pavel, N.A.; Schweizer, Th.: New Image Cleaning Method for the MAGIC Telescope. internal report TDAS-05-04, MAGIC Collaboration, 2005.

[SM06]   Schweizer, T.; MAGIC Collaboration: The MAGIC project and observations of galactic sources. In: *DPG Frühjahrstagung*. 2006.

[ST06]   Sbarufatti, B.; Treves, A.: ESO-VLT optical spectroscopy of BL Lac objects: II. New redshifts, featureless objects and classification assessments, 2006. astro-ph/0601506.

[Ste69]   Stecker, F.W.: The cosmic gamma-ray spectrum from secondary particle production in the metagalaxy. In: *The Astrophysical Journal*, volume 157, 1969.

[Ste72]   Stecker, F.W.: *Cosmic Gamma Rays.* Mono Book Corp. - Baltimore, 1972.

[Ste99]   Stecker, F.W.: Intergalactic extinction of high energy gamma-rays. In: *Astroparticle Physics*, volume 11:pp. 83–91, 1999.

[STF05]   Sbarufatti, B.; Treves, A.; Falomo, R.: Imaging Redshifts of BL Lacertae Objects. In: *The Astrophysical Journal*, volume 635(1):pp. 173–179, 2005.

[STS06]   HubbleSite. Space Telescope Science Institute (STScI), Office of Public Outreach, 2006. URL http://hubblesite.org/gallery/album/entire_collection/pr199%2027b/.

[UP95]   Urry, C.M.; Padovani, P.: Unified Schemes for Radio-Loud Active Galactic Nuclei. In: *Publications of the Astronomical society of the Pacific*, volume 107:pp. 803–845, September 1995.

[Urr00]   Urry, M.: The Hubble space telescope survey of BL Lacertae objects. II. Host glaxies. In: *The Astrophysical Journal*, volume 532:pp. 816–829, April 2000.

[Urr03]   Urry, M.: The AGN Paradigm for Radio-Loud Objects, 2003. astro-ph/0301309.

[Vas99a]   Vassiliev, V.V.: Extragalactic background light absorption signal in the 0.26-10 TeV spectra of blazars, 1999. astro-ph/9911018.

[Vas99b]   Vassiliev, V.V.: Extragalactic background light absorption signal in the TeV γ-ray spectra of blazars, 1999. astro-ph/9908088.

[Wee03]   Weekes, T.: *Very High Energy Gamma-Ray Astronomy.* IOP Publishing Ltd, 2003.

[Wei06a]   Weisstein, E.W.: Eric Weisstein's World of Physics, definition of Julian Date, 2006. URL http://scienceworld.wolfram.com/astronomy/JulianDate.html.

[Wei06b]   Weisstein, E.W.: Eric Weisstein's World of Physics (online encyclopedia), definition of Modified Julian Date, 2006. URL http://scienceworld.wolfram.com/astronomy/ModifiedJulianDate.html.

[Wit02a]   Wittek, W.: Determination of the effective observation time. internal report TDAS 02-02, MAGIC Collaboration, 2002.

[Wit02b] Wittek, W.: Image Parameters. internal report TDAS 02-03, MAGIC Collaboration, 2002.

[Wit06] Wittek, W.: private communication, 2006.

[WM05] Wagner, R.; MAGIC Collaboration: Observations of the Crab Nebula with the MAGIC telescope. In: *29th International Cosmic Ray Conference, Pune, India*, pp. 101–106. 2005.

Acknowledgement

First of all I would like to thank Prof. Dr. Nikolaj Pavel for providing the topic of this Ph.D thesis. Despite of his severe illness he always found time to give advises and to offer help with a gracious attitude. He died much too early.

I want to thankfully mention the help of Dr. Oleg Kalekin. His readiness to talk and all the many advises were invaluable. He always acted very collegial and found time to participate also in more basic work.

Furthermore I am very grateful to Dr. Wolfgang Wittek for proof-reading a first version of this thesis which lead to many improvements. He was always ready to give advises and help to solve problems.

I owe special thanks to Prof. Dr. Claus Grupen. His great interest in my work and his suggestions were very helpful.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate und Ergebnisse Anderer kenntlich gemacht habe.

Berlin, den 02.10.2006   Thomas Hengstebeck