Analysis of the Poisson Structure of H.E.S.S. Sky Maps with Minkowski Functionals

Diplomarbeit
von
Daniel Göring

Erlangen Centre for Astroparticle Physics
Physikalisches Institut
Lehrstuhl für Physik
Friedrich-Alexander-Universität
Erlangen-Nürnberg

1. Gutachter: Prof. Ch. Stegmann 2. Gutachter: Prof. K. Mecke

12. Dezember 2008

Abstract

The High Energy Stereoscopic System (H.E.S.S.) is an array of four imaging atmospheric Cherenkov telescopes, situated in the Khomas Highland in Namibia. Observing the night sky in the light of gamma-rays with energies above 100 GeV, H.E.S.S. has detected numerous new galactic and extragalactic sources of very high energy (VHE) gamma-rays. Due to its superior angular resolution, H.E.S.S. is able to resolve the spacial structure of extended sources like large shell-type supernova remnants or diffuse gamma-ray emissions from the galactic plane. This thesis introduces an analysis based on the morphology of the observed data. The structure of H.E.S.S. sky maps is quantified using the Minkowski functionals and compared to the structure of a pure Poisson background. Sources in a sky map are identified by significant deviations in the observed structure from the expected behavior.

Kurzfassung

Das High Energy Stereoscopic System (H.E.S.S.) besteht aus vier abbildenden atmosphärischen Cherenkov Teleskopen im Khomas Hochland in Namibia. Durch Beobachtungen des Nachthimmels im Lichte der Gamma-Strahlung mit Energien über 100 GeV war es H.E.S.S. möglich zahlreiche neue galaktische und extragalaktische Quellen hochenergetischer Gamma-Strahlung zu entdecken. Durch seine hervorragende Winkelauflösung ist H.E.S.S. in der Lage die räumliche Struktur ausgedehnter Quellen wie großer schalenartiger Supernova Überreste oder diffuser Gamma-Strahlung aus der galaktischen Ebene aufzulösen. In dieser Diplomarbeit wird die Analyse von Daten anhand der beobachteten Morphologie vorgestellt. Die Struktur von H.E.S.S. Himmelskarten wird durch die Minkowski Funktionale vermessen und gegen die Struktur eines reinen Poisson Hintergrunds abgeglichen. Quellen in Himmelskarten können anhand deutlicher Abweichungen in der beobachteten Struktur vom erwarteten Verhalten identifiziert werden.

Contents

1 Introduction

The discovery of cosmic rays (CR) in 1912 by Victor Hess paved the way for modern particle physics. Numerous unknown elementary particles were identified studying this radiation on mountain tops or with balloon experiments. This boom in particle physics combined with technological advances led to the development of very powerful particle accelerators, which drew the physicists from the mountains back into their laboratories. There the progress continued and led to the standard model of particle physics. Nowadays advances in accelerator-based particle physics are severely limited by the huge amount of resources needed to build and maintain these powerful tools. Therefore the CRs regain more attention, as the presence of very high energy (VHE) particles in the CRs is well established.

While today the properties of the CRs are known quite well, the origins and acceleration mechanisms are still a wide field of study. Tracking the CRs back to their sources is not a trivial task, as CRs are mainly composed of charged particles. These particles are deflected in interstellar magnetic fields and therefore loose the information about the direction of their origin. Neutral particles like gamma-rays and neutrinos travel on straight lines from their source to Earth and thus it is possible to correlate their direction on impact with the location of their origin. While the observation of cosmic neutrinos is still in an early stage of development, the field of gamma-ray astronomy has already provided remarkable results. The most successful experiment in the regime of VHE gamma-ray astronomy is the High Energy Stereoscopic System (H.E.S.S.), an imaging atmospheric Cherenkov telescope sensitive to gamma-rays with energies above 100 GeV. Due to a large field of view of 5° and its high sensitivity, H.E.S.S. is well suited for surveys of large areas in the sky.

Besides the discovery of numerous TeV point sources, the outstanding angular resolution of H.E.S.S. allowed morphological studies of extended sources like large supernova remnants for the first time in VHE gamma-ray astronomy. Thus, in addition to the correlation of VHE sources with the location of known stellar objects, it is now possible to investigate the structures of cosmic accelerators directly in the light of VHE gamma-rays. Furthermore a survey of the galactic plane revealed bands of diffuse gamma-ray emission besides many distinct sources.

The standard H.E.S.S. analysis is optimised for the search of point sources and relies on the identification of exceptionally bright spots of gamma-ray emission in the sky. The increasing amount of additional structures — like extended sources and diffuse emission — found in the H.E.S.S. observations asks for new methods for the analysis of such complex topologies. In this thesis, the analysis of H.E.S.S. sky maps based on their spacial structure is introduced. The structure of an observation is quantified using the Minkowski functionals and sources are identified by significant deviations in the observed structure from the structure of a background field.

2 Gamma-Ray Astronomy

The distribution of gamma-rays in the sky provides an impression of the most "violent" processes in the universe. In order to generate photons of such high energies, very powerful sources of particle acceleration are needed. These sources accelerate charged particles to nearly the speed of light. In secondary processes — like the decay of pions from collisions of the accelerated particles with molecules in gas clouds or inverse Compton scattering off the cosmic microwave background or ambient star light — gamma-rays are generated. As the energies of CR particles are the highest particle energies accessible to mankind, observations of these cosmic accelerators help to probe physical theories at its limits.

The detection of gamma-rays is complicated by the opacity of the Earth's atmosphere for high energy photons. Therefore, a direct observation of gamma-rays is possible only from space. While satellite-based observations provide excellent insights in low energy gamma-ray astronomy, the rapidly declining flux of CRs with increasing particle energies limits their sensitivity to at most 100 GeV. In order to be sensitive to higher energies, large detection areas are needed to account for the small flux. Using current astronautics techniques, the constraints on weight and size for an object that can be delivered to space limit the detection area of satellites to about $1\,\mathrm{m}^2$ [1].

Current ground-based gamma-ray telescopes are able to turn the atmosphere's shielding effects into a part of the detection mechanism for particle energies above 100 GeV. This permits a view on possible acceleration sites like super-massive black holes, supernova remnants and active galactic nuclei in the light of high energy gamma-rays. In addition to information on the processes in such extreme regions of our universe, the diffuse components of gamma-ray emission yield information about the density of particle accelerators in space and about the propagation of CRs in our galaxy. Therefore, the investigation of extended structures — like the diffuse emission along the galactic plane or shell-type supernova remnants — promises deep insights in the physics of high energy CRs.

2.1 Ground-based Gamma-Ray Astronomy

Even though the Earth's atmosphere is opaque for gamma-rays, they can be detected by ground-based experiments. Using secondary particles, produced by the incident gamma-rays to reconstruct the information about the primary particles, the atmosphere is turned into a part of the detector.

2.1.1 Air Showers

When a high energy photon hits the Earth's atmosphere, it interacts with the Coulomb field of an atmospheric particle and produces an electron-positron pair. These relativistic particles in turn produce high energy photons via Bremsstrahlung, which again undergo pair production. Therefore the number of electrons, positrons and photons steadily increases forming a so-called electromagnetic shower. The process continues until the mean energy per particle drops below the critical energy $E_c \approx 80\,\mathrm{MeV}$, at which point the energy loss of the shower particles due to ionisation of air molecules dominates over Bremsstrahlung and pair production. Since there are no new shower particles produced below the critical energy, the atmospheric depth at which the electrons drop below E_c is commonly referred to as the shower maximum.

In order to learn more about the characteristic parameters of an electromagnetic shower even a simple shower model helps to understand the main features. Assuming that every photon undergoes a pair production and every fermion emits a Bremsstrahlung-photon after traveling a distance of X_0 in the atmosphere, one finds the number of particles in the shower N as a function of the atmospheric depth X to be

$$N(X) = 2^{X/X_0}. \tag{1}$$

If one further assumes that in each process the energy of the primary particle is distributed equally among the two secondary particles, one finds the average particle energy E to be

$$E(X) = E_0\ 2^{-X/X_0}, \tag{2}$$

where E_0 denotes the energy of the initial gamma-ray. The shower maximum is located at

$$X_{\max} = X_0 \frac{\log[E_0/E_c]}{\log[2]} \tag{3}$$

and the corresponding maximum number of particles is

$$N_{\max} = \frac{E_0}{E_c}. \tag{4}$$

As simple as this model may be, it still describes the important shower characteristics quite well. More elaborate models and Monte Carlo simulations confirm the exponential increase of the particle number from the first interaction down to $X_{\max}$. The predicted logarithmic dependency of $X_{\max}$ on the primary energy E_0 and the proportionality of $N_{\max}$ to E_0 also hold qualitatively true for realistic air showers.

Air showers are not solely produced by gamma-rays. Actually, most of the constituents of the high energy cosmic rays interact with the atmosphere and cause air showers. Since only 1 in about 10 000 showers is induced by gamma-rays, the selection of the right kind of air showers is an important task for gamma-ray telescopes. As the majority of high energy cosmic rays are protons, the background suppression relies on the different structure of electromagnetic and hadronic air showers. While electromagnetic showers propagate as a narrow cone through the air, with only a slight broadening due to multiple scattering, hadronic showers usually exhibit a quite extended structure. Strong interactions in hadronic showers can result in secondary particles with high transverse momentum and each of these secondary particles may start its own electromagnetic sub-shower. This difference in the lateral extension as illustrated in Fig. 1 can be used to discriminate between showers caused by hadronic cosmic rays and showers due to gamma-rays (for details see Section 3.2).

2.1.2 Imaging Atmospheric Cherenkov Technique

The shower maximum of an air shower induced by a 1 TeV gamma-ray is reached about 10 km above sea level [1]. Thus, a direct detection of shower particles with ground-based experiments is only feasible for gamma-rays of very high energies. Experiments exploiting the imaging atmospheric Cherenkov technique (IACT) rely on a photograph of the air shower in the sky. This is possible, as the highly relativistic charged particles in the shower emit Cherenkov light, which can be detected in large telescopes. Cherenkov light is emitted whenever a charged particle travels faster through the ambient matter than the speed of light in this medium. This is due to the fact, that the medium is polarized by a transient

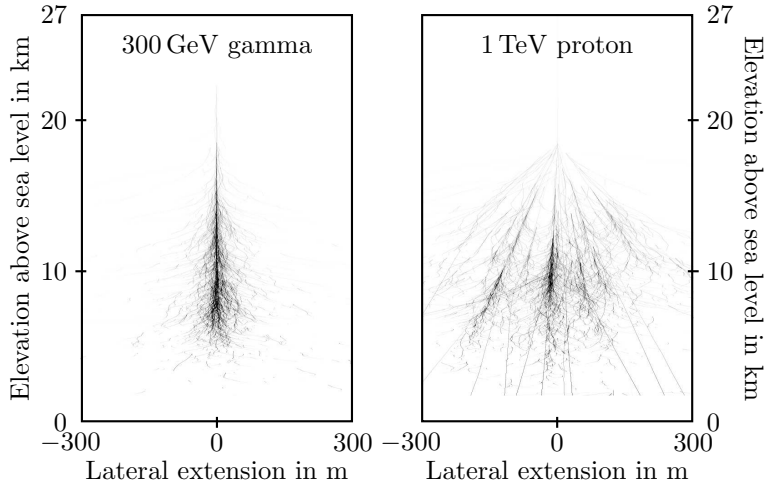

Fig. 1: *Electromagnetic (left) and hadronic (right) air showers from Monte Carlo simulations (courtesy of K. Bernlöhr). Shown are the tracks of the secondary particles.*

charged particle. All the induced dipoles behind the moving particle emit electromagnetic waves when they return to their unpolarized state. In the case of a slow particle the different dipoles interfere destructively, yielding an overall negligible effect. The contributions of the relaxing molecules behind a particle, which exceeds the speed of light in the ambient matter, on the other hand, add up to a shock front similar to a sonic boom. The Cherenkov photons are emitted in a certain angle ϑ_C to the particle trajectory. Using the notation of Fig. 2 one finds

$$\cos\vartheta_C = \frac{1}{n\beta}. \tag{5}$$

Thus, the Cherenkov angle for air shower particles changes with the refractive index of the atmosphere and with the decreasing energy along the shower axis. Taking into account the lateral extent of the shower this leads to an area of typically 250 m diameter, which is illuminated by the Cherenkov light on the ground.

As the glow from the Cherenkov light is very faint, large collectors have to be used in order to detect a reasonable amount of photons from a single air shower. Therefore IAC telescopes consist of a large mirror dish, which focuses the incident Cherenkov light onto a camera, as schematically shown in Fig. 3. While the shape of the shower is slightly changed by the optical mapping, this distortion can usually be neglected. Therefore an approximately ellipsoidal shower is mapped to an ellipse in the camera.

The IACT can be improved by stereoscopic observations. As the human eyes demonstrate impressively, a second "camera" — even if quite close to the first one — provides a perception of depth and can be used to reconstruct the three-dimensional orientation of objects in space. Thus, by placing a second IAC telescope within the light pool of the same shower, the direction reconstruction can be improved significantly. In addition, a second observer helps to significantly reduce the number of background events. On the one hand, this is simply due to the additional information about the shower, which helps to determine its primary particle. On the other hand, monoscopic observations suffer from an additional background due to the Cherenkov light of muons. As the light cone of muons is very small, muonic events are only seen in one camera and can be rejected in stereoscopic observations.

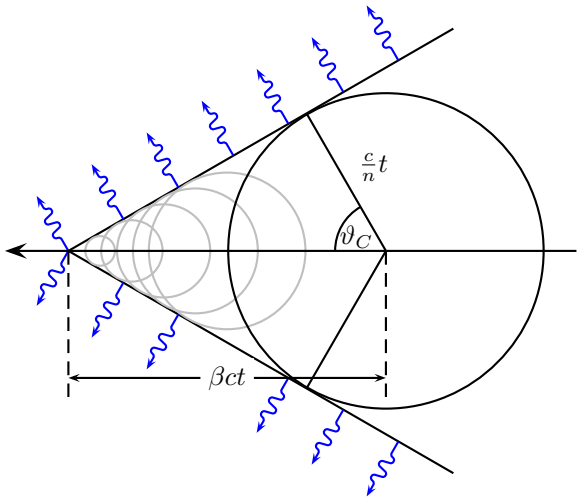

Fig. 2: *Formation of a Cherenkov cone for a charged particle, which travels with a velocity of $\beta c > c/n$ through a medium with refractive index n. The circles indicate spherical electromagnetic waves emitted by the relaxation of induced dipoles. The contribution of these waves along the path of the charged particle form a shock front with a characteristic opening angle.*

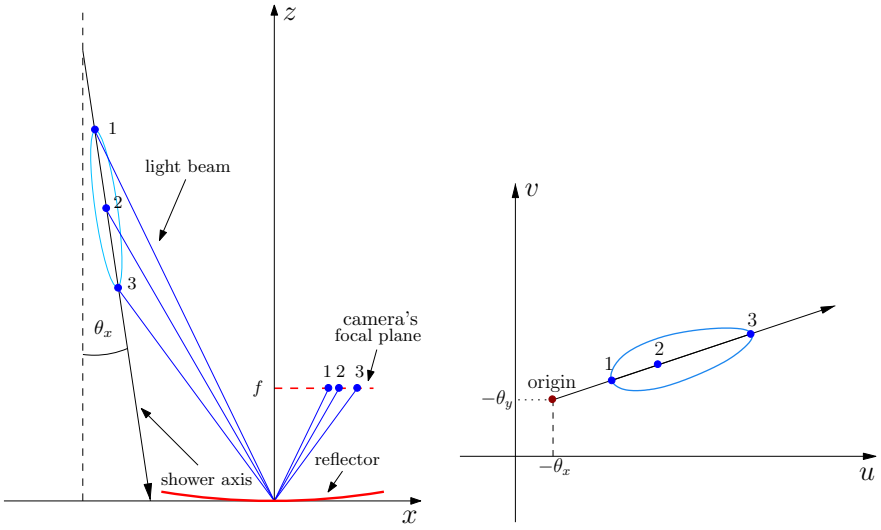

Fig. 3: *Schematic view of an imaging atmospheric Cherenkov telescope. The Cherenkov light (blue) is collected by a large reflector dish (red) and focused onto a camera. The resulting camera image is illustrated on the right.*

3 The H.E.S.S. Experiment

The High Energy Stereoscopic System (H.E.S.S.) is at the moment the most advanced IACT experiment. It is situated in the Khomas Highland in Namibia, about 100 km south-west of the capital Windhoek. The H.E.S.S. site with its four Cherenkov telescopes is shown in Fig. 4. The location in the south of Africa has been chosen for its excellent observation conditions. Especially the many cloudless nights and virtually no light pollution are vital conditions for IACT experiments. Using the stereoscopic approach and large collector dishes, H.E.S.S. achieves a good background rejection and high sensitivity for gamma-rays above 100 GeV. This sensitivity allows the detection of gamma-ray sources with a flux of 1% of the flux of the Crab nebula within 25 h of observation. Single gamma-ray events can be reconstructed with an angular resolution of a few arcseconds. With its big field of view of about 5°, H.E.S.S. can observe an area of approximately 10 times the diameter of the full moon on the night sky at once. These properties lead to the discovery of numerous new TeV sources since the commencement of operation.

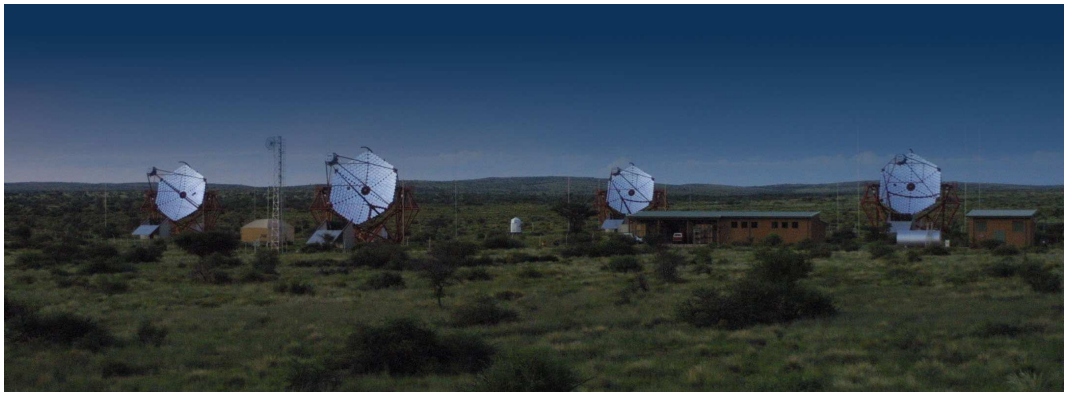

Fig. 4: *The four H.E.S.S. telescopes in the Khomas Highland in Namibia.*

3.1 The H.E.S.S. Telescopes

The four IAC telescopes of the H.E.S.S. experiment are arranged in a square of side length 120 m. This distance is chosen, as it corresponds to the typical radius of the Cherenkov light footprint of a gamma-ray shower. Thus, showers can usually be seen by at least two telescopes, while the angle between the different observing telescopes is maximised, yielding the best stereoscopic observation.

Each of the four telescopes consists of a big mirror dish on an alt-azimuth mount, which focuses the collected Cherenkov light onto a camera (see Fig. 5). The reflector dish consists of 382 round mirrors of 60 cm diameter arranged in a Davies-Cotton design on a sphere of of radius 15 m around the camera, which provides good imaging also for events further away from the optical axis. This setup yields a total mirror area of 108 m^2 per telescope. Each of the mirrors can be adjusted by two remote-controlled motors, for automatic realignment.

The reflected light is detected by a camera of 960 pixels. These pixels are arranged in a hexagonal grid, where each hexagon is connected to a photomultiplier tube (PMT) via a Winston cone, which guides the incident light to the sensitive area of the PMT. Since the

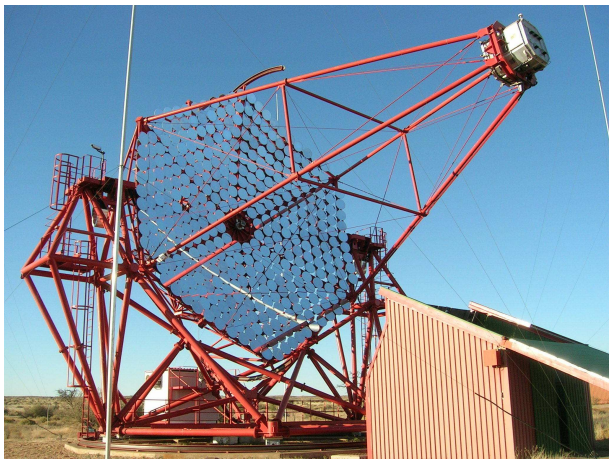

Fig. 5: *One of the imaging atmospheric Cherenkov telescopes of the H.E.S.S. experiment.*

flash of Cherenkov light from an air shower lasts only a few nanoseconds, the camera needs to have fast electronics in order to resolve single gamma-ray induced showers. Therefore, the signal of each PMT is sampled with a rate of 1 GHz. As this enormous amount of data would put too much load on the connected data acquisition system (DAQ), the data is reduced by a two level trigger system. The first level is processed directly in the camera electronics by the so called camera trigger. For this, the camera is divided into 38 overlapping trigger sectors of 8×8 pixels. A configurable amount of pixels within one trigger sector with a signal above a certain threshold during a coincidence window of about 1.5 ns leads to a camera trigger. Camera triggers are sent from the cameras to the second level trigger in the control building, the so-called central trigger. As H.E.S.S. is intended for the stereoscopic observation of air showers, the central trigger checks for coincidences in the camera triggers of different cameras and initiates the read-out of the camera buffers only if the coincidence requirements are met. These trigger conditions help to reduce the recorded background and therefore improve the sensitivity and lower the energy threshold of the experiment.

3.2 Event Reconstruction

Each H.E.S.S. observation is divided in runs of about 28 min. These runs have to pass data quality checks before further analysis to rule out hardware problems or bad observation conditions like cloudy skies. Successful runs contain camera images of $\mathcal{O}(10^3)$ air showers, from which the primary gamma-rays need to be reconstructed. In a first step, a so-called image cleaning is applied, which is supposed to isolate the shower in the image and remove the ambient night sky background. Therefore, only clusters of high-intensity pixels are kept in the image, that is pixels of at least 10 photoelectrons (p.e.) are kept, if at least one of the neighbouring pixels has an intensity of at least 5 p.e. and pixels of at least 5 p.e. are kept, if an adjacent pixel of at least 10 p.e. exists. The remaining structure is approximated by an ellipse and parametrized by the Hillas parameters [2]. These parameters are shown in Fig. 6 and include the *length* (length of the semimajor axis), the *width* (length of the semiminor axis), the *distance* (distance of the image's center of gravity to the camera center), the angle α (orientation of the ellipse) and the *size* (total number of p.e. in the image). The Hillas parameters are used to reconstruct the direction of the shower axis and thus the direction

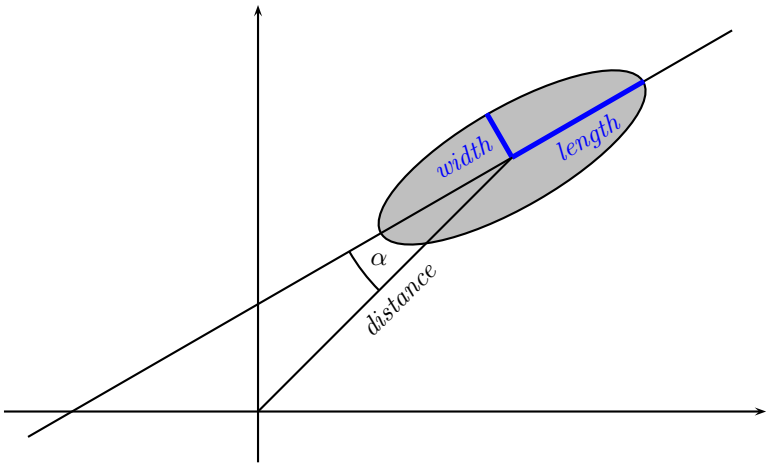

Fig. 6: *Hillas parameters describing shape and position of a camera image.*

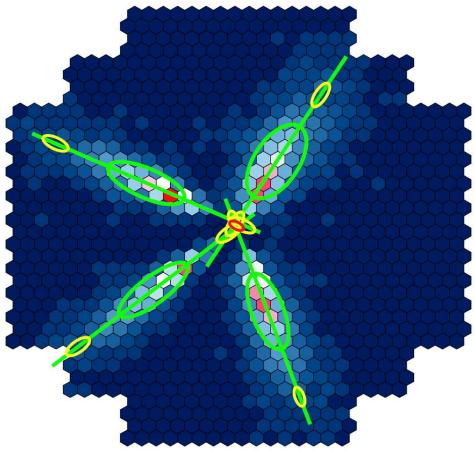

Fig. 7: *Reconstruction of the shower origin for a* 7 TeV *gamma-ray shower seen by all four telescopes. The four images from the different cameras are superimposed in a common view. The Hillas ellipses for each of the single images and their major axes are overlaid in green. The yellow areas show an estimate of the shower origin based on the single cameras images, while the red region gives the estimate of the origin by the stereoscopic reconstruction. (Figure taken from [3])*

of the initial gamma-ray. In order to do this, the origin of the shower and its impact point on the ground are determined. As the four telescopes observe the same region of the sky, the origin of the shower has to be in the same place in every camera image of the night sky. Therefore the shower origin can be reconstructed by superimposing several camera images from a stereoscopic observation and determining the common origin of the showers in the combined image (see Fig. 7). Since the intersections of the major axes of every pair of camera images may yield slightly different origins, the pairwise intersections are weighed according to the uncertainties in the involved camera images to determine the best overall estimate [4]. The point of impact is determined similarly. Instead of projecting the shower images onto the sky, they are projected onto the ground. In this reference frame the camera centers do not reference a common point, but are projected onto the according telescope's position. The intersections of the images' major axes in this projection are also weighed according to the uncertainties in the involved images to find the best estimate for the impact location.

As shown in Fig. 1, the shape of an air shower can be used to discriminate between hadronic cosmic rays and gamma-rays. Since the shape of each camera image of a shower is given by the Hillas parameters *length* and *width*, it is possible to use the Hillas parameters for gamma-hadron separation. The expected shape for an image of a gamma-ray shower depends on the primary energy of the incident particle and the distance of the shower to the observing telescope. Knowing the origin and point of impact of the shower (see above), the distance of the shower to each telescope (so-called impact parameter) can be calculated. The amount of light deposited by a shower (*size*) with a certain impact parameter gives an estimate of the energy of the primary gamma-ray. This estimate depends on the observation position and varies with the zenith angle[1] of the observed target. The expected shape for given impact parameter, *size* and zenith angle is determined by Monte Carlo simulations and used to calculate the reduced scaled width and length:

$$\begin{aligned} \mathrm{RSW} &= \frac{width - \langle width \rangle}{\sigma(width)} \\ \mathrm{RSL} &= \frac{length - \langle length \rangle}{\sigma(length)} \end{aligned} \tag{6}$$

The reduced scaled parameters of the different telescopes are weighed and averaged to determine the shape parameters of the observed shower, the so-called mean reduced scaled width and length.

$$\mathrm{MRSW} = \frac{\sum_{i=1}^{N} w_i \mathrm{RSW}_i}{\sum_{i=1}^{N} w_i}, \tag{7}$$

where RSW_i is the reduced scaled width of the i^{th} telescope, N the number of telescopes and

$$w_i = \frac{\langle width_i \rangle^2}{\sigma^2(width_i)} \tag{8}$$

the weight of the i^{th} telescope. According definitions hold for the mean reduced scaled length MRSL. The optimal selection criteria for the MRSW and MRSL parameters depend on the intensity and energy spectrum of the observed gamma-ray source and were determined for several source classes by extensive Monte Carlo simulations. To minimize an a priori bias on the results presented in this thesis, only the H.E.S.S. standard cuts [5] are used in this

[1] Angle between the observation position and a vertical axis pointing towards the sky.

work. These cuts select events with $-2.0 \leq \mathrm{MRSL} \leq 2.0$ and $-2.0 \leq \mathrm{MRSW} \leq 0.9$. Due to this event selection, the signal to background ratio for strong point sources is reduced from about 1 : 100 for triggered events to approximately 1 : 1 after cuts.

3.3 Data Analysis

The background remaining after all selection cuts has to be estimated, in order to be able to identify signals of gamma-ray sources. This background consists mainly of hadronic events with an air shower shaped similarly to the shape of a gamma-ray shower. Additionally, there are events from diffuse gamma-rays, which are created by high energy cosmic rays interacting with interstellar gas, photons of starlight or the cosmic microwave background. The amount of background events is estimated by dedicated background measurements. To be comparable to the initial observation, these so-called off-runs need to be taken under identical observation conditions. Usually, nearby regions (off-regions) are used to estimate the background flux. This can be done using dedicated off-runs, but due to the large field of view of H.E.S.S. it is also possible to compare different regions in the field of view of the same run. Several different methods of choosing off-regions and estimating the background from these regions are discussed in [6]. The different background models have different advantages and disadvantages for different source types or properties under study, but all of them yield compatible results. In this thesis, only the ring background model will be introduced, as it is very robust and usually the preferred method at the start of an analysis.

Given a photon count N_{on} in the observation region (on-region) and a number of N_{off} in the off-region, the excess is given by

$$N_{\mathrm{ex}} = N_{\mathrm{on}} - \alpha N_{\mathrm{off}}, \tag{9}$$

where α is the ratio of effective exposure integrated in time and angular space over the on- and off-region[2]. Under the hypothesis of Poisson distributed background, the significance of an excess can be determined by a likelihood ratio estimator [7]. The resulting significance

$$S = \sqrt{2\left\{N_{\mathrm{on}} \ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{\mathrm{on}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right] + N_{\mathrm{off}} \ln\left[(1+\alpha)\left(\frac{N_{\mathrm{off}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right]\right\}} \tag{10}$$

follows a standard normal distribution, if no additional source is present in the on-region. Values for α and N_{off} have to be supplied by the background model.

The name of the ring background model is derived from the way the on- and off-regions are chosen from the field of view. The excess for a certain position in the field of view is determined by defining the on-region to be a disc centered at the test position. The radius of the disc is chosen according to the experiment's angular resolution of a few arcminutes in order to collect all events from a possible point source at the test position. An annular region around the test position is chosen as off-region (see Fig. 8). The area of the ring is usually about 7 times the area of the on-region, but this does not immediately yield $\alpha \approx 1/7$, due to different observation conditions in the ring and in the on-region. These differences originate in a spacial variability of the camera acceptance. Assuming a radial symmetric structure of the sensitivity, the acceptance can be modeled from the measured data. In order to do this, the number of events as a function of the angular distance squared ψ^2 to the observation

[2] For identical observation conditions in on- and off-region, one finds $\alpha = 1$. For e.g. an off-region of twice the size as the on-region or an off-run twice as long as the on-run and otherwise identical conditions, one finds $\alpha = 1/2$.

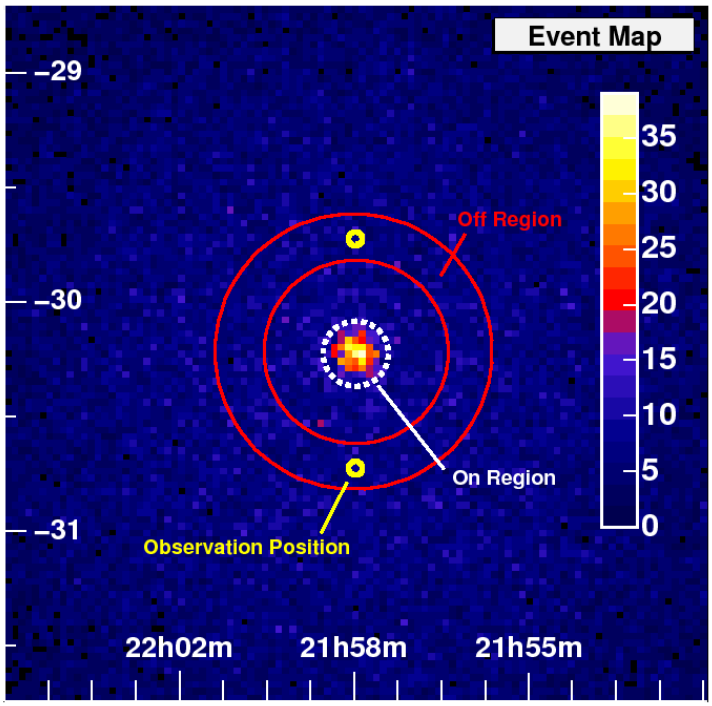

Fig. 8: *Exemplary view of an on- and off-region of the ring background model [6].*

position in the center of the field of view is determined. For a pure off-run without any sources in the field of view, this is done by summing the event counts in concentric rings around the observation position and dividing each count by the area of its ring. In the case of known sources in the field of view, all rings close to the source are segmented and segments next to the source are ignored; the event count is divided by the area of the valid segments. The radial distribution found this way is fitted by a 5^{th} degree polynomial to determine the final acceptance model. Knowing the camera acceptance, the field of view can be corrected for the acceptance, in order to equalize the observation conditions for different regions on the observed sky. Thus, the sky map is acceptance corrected by dividing each bin's event count by the corresponding camera acceptance of its position. For the acceptance-corrected sky map α_c is simply given by

$$\alpha_c = \frac{A_{\text{on}}}{A_{\text{off}}} \approx \frac{1}{7}, \tag{11}$$

with $A_{\text{on/off}}$ being the area of the on- respectively off-region. The excess photon count for a certain region in the field of view can now be determined by the acceptance-corrected event count of this region ($N_{\text{acc,on}}$), the acceptance corrected event count in the ring around the on-region ($N_{\text{acc,off}}$) and α_c. The significance of the obtained excess is determined using Eq. (10).

3.4 Observed Structures

Using the presented strategy, the H.E.S.S. experiment discovered more than 40 new TeV sources during its first four years of operation. An impressive step forward, bearing in mind that less than 10 sources were established before the commencement of H.E.S.S. Among these new sources one finds several extragalactic and galactic point sources. However, many of the discovered galactic sources seem to be extended (see Fig. 9). Due to the good angular resolution of H.E.S.S., the morphology of objects like the large shell type supernova remnant RX J1713.7 − 3946 (galactic longitude 347.33°, galactic latitude −0.47°; cp. Fig. 9) could be studied in the regime of VHE gamma-rays for the first time [8].

In addition to the structure of extended sources, a global structure is apparent in the view of the galactic plane: The galactic sources are perfectly aligned along the plane and cluster towards the Galactic Center. This leads to observations with several sources in the same field of view. Besides such discrete regions of gamma-ray emission, an extended band of diffuse gamma-ray emission was found close to the Galactic Center [9]. Complex configurations of multiple sources during one observation, extended sources and the diffuse gamma-ray emission along the galactic plane raise the question for an analysis, which takes the structure of the observed data into account.

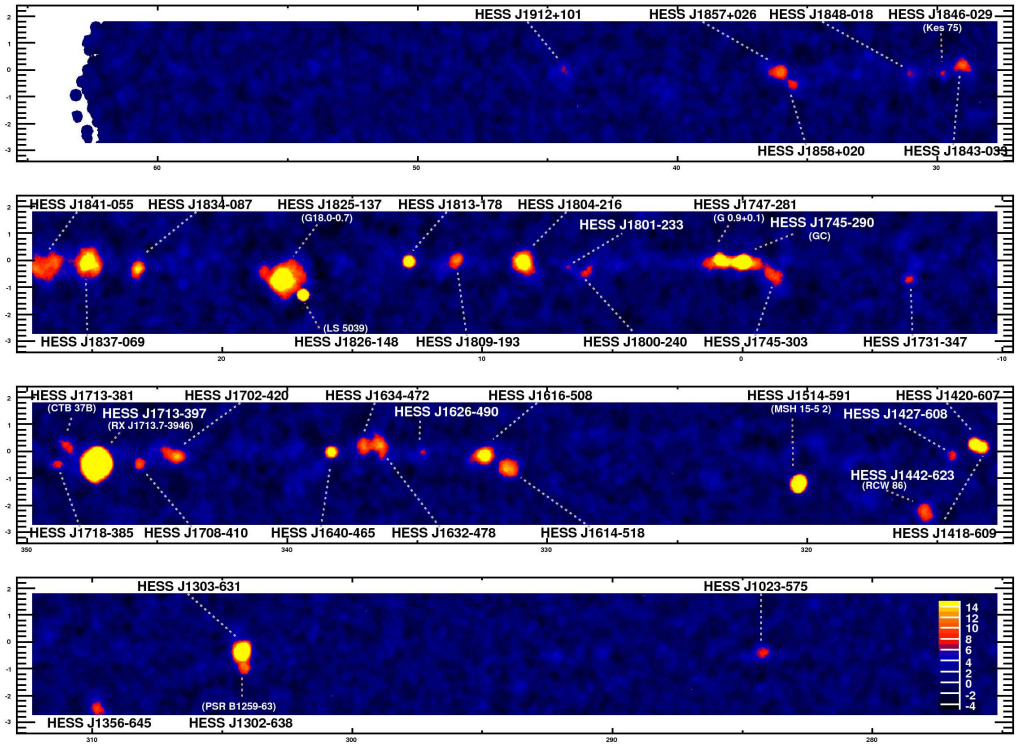

Fig. 9: *Galactic plane as seen by H.E.S.S. in 2007. The color indicates the significance of the observed signal in each bin and the position is given in galactic coordinates.*

4 How to Quantify Structure?

The human eye is usually very good in finding patterns and recognizing structures. Unfortunately, it is not obvious how to translate these subjective perceptions into measurable quantities. But for a reliable and verifiable analysis of structures, well-defined measures are vital.

In a first approach, one may try to find suitable measures in the mathematical sense [10]. What properties should such a measure have? As one is interested in deviations from a homogeneous isotropic background, the measure should be invariant under translations and rotations. But in $\mathbb{R}^n$ there is — up to a positive multiplicative constant — only one motion-invariant measure: The Lebesgue measure. Thus, confining oneself to measures in the mathematical sense, one ends up with the volume of a body as the only quantifiable information about its structure.

As the volume certainly gives some information on the morphology of a body, the properties of a measure may be considered the right starting point. Before trying to find further quantities by giving up some of the constraints a mathematical measure has to fulfill, an additional constraint is imposed, which proofs itself to be very helpful and not very constraining: Instead of dealing with the whole set of measurable sets, one should concentrate on the so called convex ring $\mathcal{R}$. The convex ring comprises all compact convex sets and all finite unions of compact convex sets — the so-called poly-convex bodies. As physical data is usually binned and pictures are pixelated (voxelated), the restriction to $\mathcal{R}$ is naturally met; all bodies are composed of convex pixels (voxels) and therefore lie within $\mathcal{R}$. Abandoning the positiveness of the measure and changing the σ-additivity to additivity while keeping a convex continuity, one ends up with a useful class of "measures", the so called quermass integrals or Minkowski functionals. These functionals and their properties are presented in the following section.

4.1 Minkowski Functionals

A functional $\varphi : \mathcal{R} \to \mathbb{R}$ is called additive, if

$$\varphi(K_1 \cup K_2) = \varphi(K_1) + \varphi(K_2) - \varphi(K_1 \cap K_2) \quad \forall\, K_1, K_2 \in \mathcal{R}. \tag{12}$$

As it is in general not desired to count features of an image more than once, additivity is a natural requirement for image functionals.

Moreover, the quantification of a structure should be independent of its orientation or position in the image. Therefore, another reasonable constraint for image functionals is motion-invariance:

$$\varphi(mK) = \varphi(K) \quad \forall\, K \in \mathcal{R}, m \in \mathcal{M}, \tag{13}$$

where $\mathcal{M}$ denotes the group of motions (translations and rotations) in the Euclidean space.

As a final constraint, the functional of a convex body may be approximated by the functional of a "similar" convex body. In mathematical terms

$$\lim_{i \to \infty} \varphi(K_i) = \varphi(K), \tag{14}$$

where $K_i, K \in \mathbb{K}$, with $\mathbb{K} \subset \mathcal{R}$ being the space of compact convex bodies, and $K_i \xrightarrow{i\to\infty} K$ in the Hausdorff metric on $\mathbb{K}$. As this continuity is required only on $\mathbb{K}$ and not on the whole convex ring $\mathbb{R}$, it is called convex continuity.

All these properties are fulfilled by the Minkowski functionals. Moreover a remarkable theorem by Hadwiger [11] states that every additive, motion-invariant and convex continuous functional on $\mathcal{R}$ can be written as

$$\varphi(K) = \sum_{i=0}^{d} a_i W_i^{(d)}(K), \tag{15}$$

where $W_i^{(d)} : \mathcal{R} \to \mathbb{R}$ are the Minkowski functionals and $a_i \in \mathbb{R}$ are suitable coefficients. Furthermore, the formula implies that there are only $d+1$ independent functionals for a d-dimensional body.

It lies beyond the scope of this thesis to properly introduce and define all Minkowski functionals for a general d-dimensional body, but fortunately there are intuitive geometric interpretations of the functionals in $\mathbb{R}^2$. The Minkowski functionals are proportional to the area A, perimeter P and Euler characteristic[3] χ of a body [12]:

$$W_0^{(2)} \propto A \qquad W_1^{(2)} \propto P \qquad W_2^{(2)} \propto \chi \tag{16}$$

The introduced formalism is well-suited to quantify the structure of black and white images. In order to investigate more complex data like grey scale images or intensity distributions, one usually uses excursion sets and analyzes their morphology. An excursion set $\mathcal{D}_\varrho$ of an intensity distribution $I : \mathbb{R}^2 \to \mathbb{R}$ is defined as

$$\mathcal{D}_\varrho = \{\vec{x} \in \mathbb{R}^2 | I(\vec{x}) \geq \varrho\}. \tag{17}$$

By interpreting the excursion set as the black regions of an image and the remaining parts as white regions, the data is mapped to a threshold dependent black and white image. Since the shape of an excursion set depends on the threshold parameter ϱ, the values of the Minkowski functionals also depend on ϱ. Thus one ends up with three functions describing the structure of the intensity distribution: $A(\varrho), P(\varrho)$ and $\chi(\varrho)$.

While the evaluation of the Minkowski functionals on $\mathbb{R}^2$ is quite complex and involves nontrivial integrals, it is very easy for a grid of black and white square pixels. Instead of looking at the individual pixels of the image, one decomposes the image even further into vertices, open edges and the open interior of each pixel (see Fig. 10). Due to the additivity of the Minkowski functionals, the plain number of black interiors, edges and vertices is enough to determine the functionals. All in all the evaluation reduces to a simple counting of black picture elements. Be n_0 the number of black vertices, n_1 the number of black open edges and n_2 the number of black interiors, then the Minkowski functionals can be written in an amazingly simple way [12]:

$$A = n_2 \qquad U = 2n_1 - 4n_2 \qquad \chi = n_0 - n_1 + n_2. \tag{18}$$

Using this definition, the Minkowski functionals can be applied efficiently in the structure analysis of two-dimensional data.

[3] The Euler characteristic is defined to be $\chi(K) = 1 \quad \forall\, K \in \mathbb{K}$ and can be determined for all bodies in $\mathcal{R}$ by decomposing them into compact convex bodies and using the additivity of χ to sum up the contributions. As the Euler characteristic of an image with n disjoint convex bodies is n, χ is often called the connectivity number.

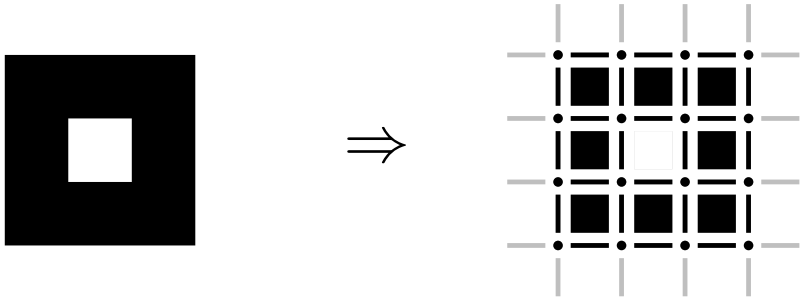

Fig. 10: *Decomposition of a b/w pixel image into vertices, open edges and open interiors of pixels.*

4.2 Application to Poisson Fields

In order to apply the Minkowski functionals to the analysis of H.E.S.S. sky maps, the typical structure of a background measurement has to be known. Additional structures besides the background — like gamma-ray sources — can then be identified by deviations in the observed structure from the expected behavior.

A common background model is the hypothesis of Poisson distributed background events. Additionally, one usually assumes the spatial distribution of the events to be featureless, thus the structure to be homogeneous and isotropic. In this section the application of the Minkowski functionals to such a homogeneous and isotropic Poisson field is presented.

4.2.1 Expected Structure of Poisson Fields

A Poisson field is a random function $\psi : \mathbb{Z}^2 \to \mathbb{N}_0$, where $\psi(\vec{x})$ follows a Poisson distribution $\forall\ \vec{x} \in \mathbb{Z}^2$. To meet the constraints of a homogeneous and isotropic field, the number of expected events is required to be constant throughout the field $\langle \psi(\vec{x}) \rangle = \lambda \quad \forall\ \vec{x} \in \mathbb{Z}^2$. Furthermore, $\psi(\vec{x}_1)$ and $\psi(\vec{x}_2)$ are statistical independent $\forall\ \vec{x}_1 \neq \vec{x}_2 \quad \vec{x}_1, \vec{x}_2 \in \mathbb{Z}^2$ for pure background.

Physical data can be mapped to such a model by identifying the center of each bin of a two-dimensional count map with one point in $\mathbb{Z}^2$. However, as real detectors have a finite sensitive area, one never has a field over the whole $\mathbb{Z}^2$, but only over a finite sub region. To minimize boundary effects, the physical field is extended to $\mathbb{Z}^2$ by imposing periodic boundary conditions[4]. Nevertheless, the Minkowski functionals are always determined in the finite sub region, as the periodic extension would just increase the value of the functionals by a factor equal to the number of additional areas (see Eq. (12)).

$A(\varrho), P(\varrho)$ and $\chi(\varrho)$ can be easily determined for the Poisson fields. The probability of one pixel to be above the threshold ϱ is given by

$$p(\varrho) = \sum_{k=\lceil \varrho \rceil}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}, \tag{19}$$

[4]Each row and each column of pixels is extended by an identical row or column of pixels.

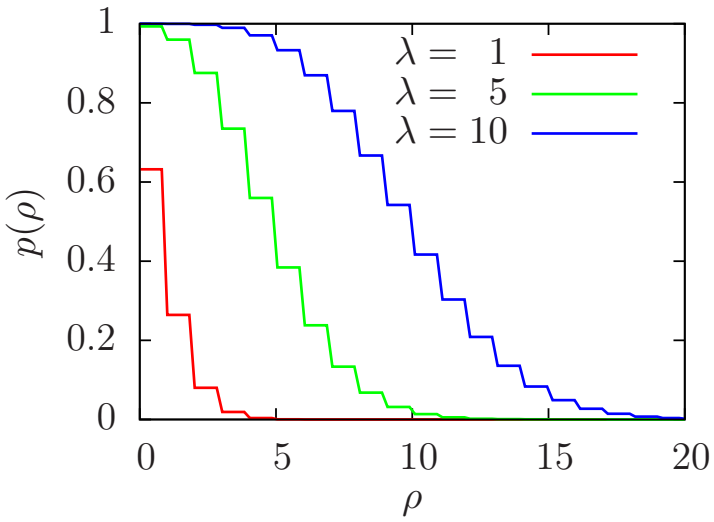

Fig. 11: *Probability $p(\varrho)$ of a pixel to be black from Eq. (19) for different values of λ.*

which is shown in Fig. 11 for different values of λ. Then the probability of a black edge is given by

$$\begin{aligned} p_1(\varrho) &= 1 - (1 - p(\varrho))^2 \\ &= p(\varrho)(2 - p(\varrho)). \end{aligned} \tag{20}$$

This follows from the fact that an edge is only counted as white, if both adjacent pixels are white, and the probability for this complementary event is given by $(1 - p(\varrho))^2$, as the neighbouring pixels are independent. Similarly, the probability of a black vertex is given by

$$\begin{aligned} p_0(\varrho) &= 1 - (1 - p(\varrho))^4 \\ &= p(\varrho)(2 - p(\varrho))(2 - 2p(\varrho) + p^2(\varrho)). \end{aligned} \tag{21}$$

Hence, the expected numbers of black pixels $\langle n_2 \rangle$, black edges $\langle n_1 \rangle$ and black vertices $\langle n_0 \rangle$ for a field of $N \times N$ bins are given by

$$\begin{aligned} \langle n_2 \rangle &= N^2 p(\varrho) \\ \langle n_1 \rangle &= 2N^2 p_1(\varrho) \\ \langle n_0 \rangle &= N^2 p_0(\varrho). \end{aligned} \tag{22}$$

Using Eq. (18) one finally finds

$$\begin{aligned} \langle A \rangle(\varrho) &= N^2 p(\varrho) \\ \langle P \rangle(\varrho) &= 4N^2 p(\varrho)(1 - p(\varrho)) \\ \langle \chi \rangle(\varrho) &= N^2 p(\varrho)(1 - p(\varrho))(1 - 3p(\varrho) + p^2(\varrho)). \end{aligned} \tag{23}$$

The general shapes of the expectation values of the Minkowski functionals are shown in Fig. 12.

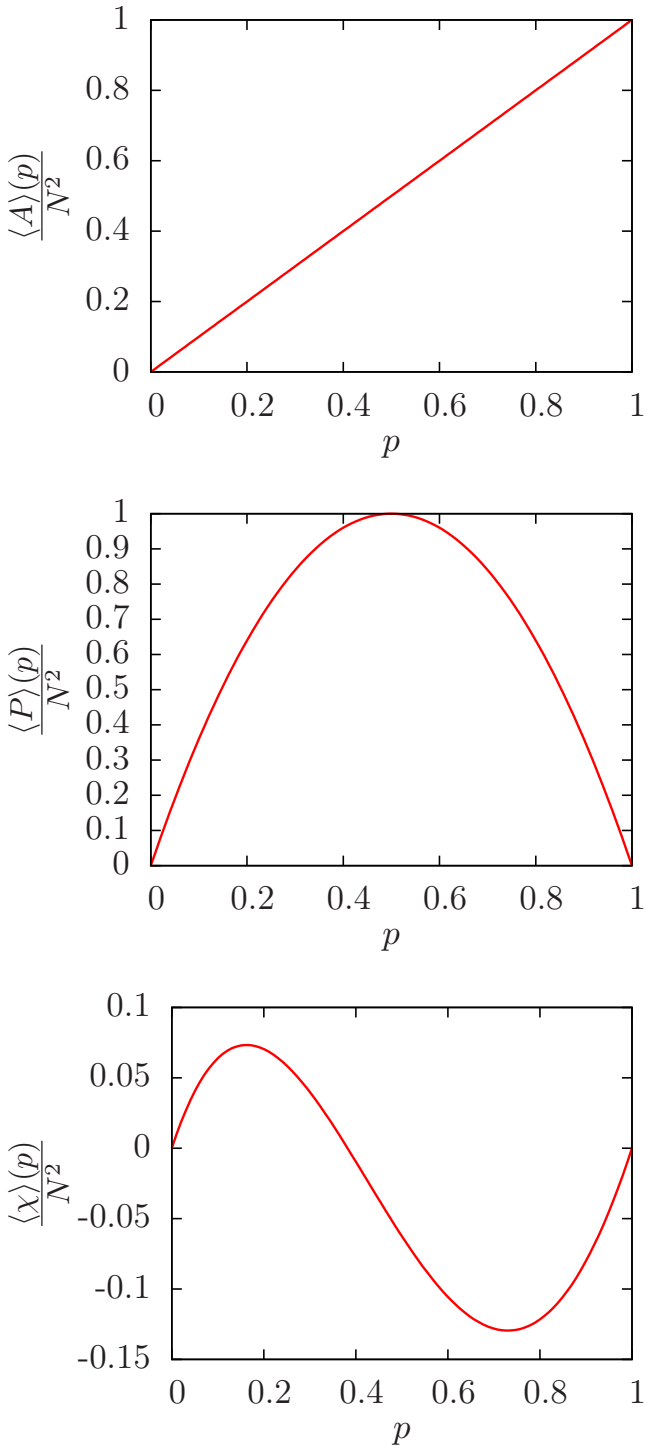

Fig. 12: *The expectation values of the Minkowski functionals A, P and χ as a function of the probability p of a pixel to be black.*

4.2.2 Quantifying Fluctuations of Poisson Fields

So far, the expectation values of the Minkowski functionals have been derived. However, in reality every background measurement exhibits fluctuations around the expected structure. In order to be able to discriminate between statistical fluctuations and significant deviations from the background structure, it is vital to quantify deviations from the expected morphology. At first this problem is investigated in detail for the area $A(\varrho)$.

As all pixels are independent, $A(\varrho)$ follows a binomial distribution

$$\Pr(A(\varrho) = k) = \binom{N^2}{k} p^k(\varrho)(1 - p(\varrho))^{N^2-k}. \tag{24}$$

Since the distribution is highly asymmetric for very low and very high thresholds, it is not sufficient to compare a deviation from the expected value to the standard deviation[5]. Instead, one has to calculate the probability of finding an according fluctuation in a pure background field. The compatibility of a measured value k with the distribution of a general random number X be defined as:

$$\mathrm{comp}(X = k) = \begin{cases} \Pr(X \leq k) & \forall\, k \leq \langle X \rangle \\ \Pr(X \geq k) & \forall\, k > \langle X \rangle \end{cases} \tag{25}$$

Thus, the compatibility of k to the distribution of X is the probability to find a statistical fluctuation in X, which is at least as far from the expected value $\langle X \rangle$ as k. Therefor, the compatibility of a measured area A_m at a certain threshold ϱ with the Poisson field $\mathrm{comp}(A(\varrho) = A_m)$ is the sum of the probabilities to measure a value of A_m or an even stronger deviation in a pure background field. While the distribution of $A(\varrho)$ is known and deviations can be quantified accordingly, the quantification of deviations for the remaining Minkowski functionals is more challenging.

Quantifying Fluctuations of small Poisson Fields For small fields, one can simply count all possible field configurations and deduce the according distributions by "brute force". But as the number of field configurations for a $N \times N$ field is 2^{N^2}, the required time increases vastly. For a 4×4, field an ordinary 2 GHz PC can construct all distributions in less than 1 s and a 5×5 field is handled in less than 1 min. Beyond this size, the over-exponential complexity prevails and a 6×6 field takes already about 2 days to be processed; extrapolating this development to a 7×7 field one can expect the program to run for more than 30 years to process all $2^{49} \approx 10^{15}$ configurations.

Even though an image of 6×6 pixels may seem very small, it is actually big enough to reveal some of the general features of the desired distributions. The variances of the Minkowski functionals have the same form for $N = 3$, $N = 4$, $N = 5$ and $N = 6$. One finds

$$\begin{aligned} \mathrm{Var}[A](\varrho) &= N^2 p(\varrho)(1 - p(\varrho)) \\ \mathrm{Var}[P](\varrho) &= 8N^2 p(\varrho)(1 - p(\varrho))(2 - 7p(\varrho) + 7p^2(\varrho)) \\ \mathrm{Var}[\chi](\varrho) &= N^2 p(\varrho)(1 - p(\varrho)) \cdot \\ &\qquad (1 - 12p(\varrho) + 64p^2(\varrho) - 139p^3(\varrho) + 137p^4(\varrho) - 59p^5(\varrho) + 9p^6(\varrho)). \end{aligned} \tag{26}$$

[5] This corresponds to a Gaussian approximation of the distribution, which is inadequate for distributions with a high skewness.

Likewise, the variances of the number of black pixels, edges and vertices exhibit the same form for the given field sizes.

$$\begin{aligned}
\mathrm{Var}[n_2](\varrho) &= N^2 p(\varrho)(1-p(\varrho)) \\
\mathrm{Var}[n_1](\varrho) &= 2N^2 p(\varrho)(1-p(\varrho))^2(8-7p(\varrho)) \\
\mathrm{Var}[n_0](\varrho) &= N^2 p(\varrho)(1-p(\varrho))^4(16-38p(\varrho)+32p^2(\varrho)-9p^3(\varrho))
\end{aligned} \tag{27}$$

For n_i it is shown in appendix A that the given formulas for the variances are actually valid for all $N > 4$. From these results it is very plausible to infer the validity of Eq. (26) for large N.

Even though the distributions for $N = 6$ are available, there are technical reasons to prefer $N = 5$ as reference. With 2^{49} different configurations the 6×6 field has more states than an 32 bit integer can count. Indeed, there are even coefficients in the distributions for $N = 6$, which are larger than $2^{32} - 1$. As the use of 64 bit architectures and especially operating systems is not yet very common, the use of 64 bit integers to implement the distributions would lead to a severe speed penalty and to compatibility issues. Therefore, the distributions of the 5×5 field are chosen over the distributions for $N = 6$ throughout this work. The distributions of the Minkowski functionals for $N = 5$ can be seen in Fig. 13. Knowing the exact distributions for small fields, one can use them to compute the probability for a deviation to be compatible with a random fluctuation as introduced in Eq. (25).

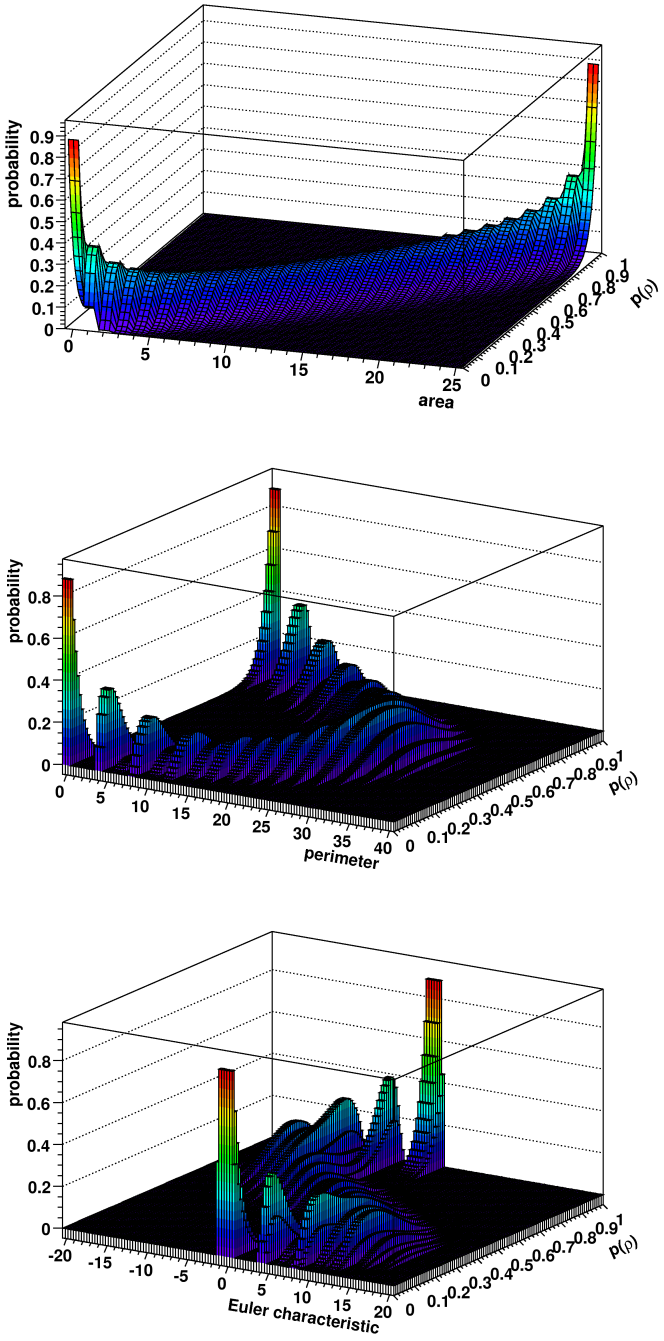

Fig. 13: *The distributions of the Minkowski functionals A, P and χ for an image of 5×5 pixels as a function of the probability p of a pixel to be black.*

Quantifying Fluctuations of large Poisson Fields In the case of large fields, one has to switch to approximations to determine the compatibility of a given configuration with the background model. To find an approximation for the perimeter P it is helpful to use its original geometric definition (instead of the definition given in Eq. (18)). Thus, the perimeter of the field is equal to the number of edges with different colors on each side. Since the probability of finding two pixels with different color next to each other is $2p(\varrho)(1-p(\varrho))$ (see Fig. 14) and there are $2N^2$ edges in a $N \times N$ field, one finds

$$\langle P \rangle(\varrho) = 4N^2 p(\varrho)(1 - p(\varrho)), \tag{28}$$

consistent with Eq. (23). As the pixels are statistically independent, each edge is only correlated with six neighbouring edges (see Fig. 15). Therefore, the correlations are confined within small areas and can be neglected for large fields. But there is one geometric constraint, which must not be neglected: P has to be an even number. For each edge separating a black and a white region, an associated edge can be found that detaches the respective region at the opposite side[6]. Thus, $2p(\varrho)(1 - p(\varrho))$ is actually the probability to find *two* perimeter sections and there are N^2 pairs of edges in the image. Therefore, the following binomial distribution can be used to approximate the distribution of the perimeter for large fields:

$$\Pr(P(\varrho) = 2k) \approx \binom{N^2}{k} (2p(\varrho)(1 - p(\varrho)))^k (1 - 2p(\varrho)(1 - p(\varrho)))^{N^2-k} \tag{29}$$

A comparison of this distribution with the exact one for $N = 5$ shows a similar structure but discrepancies especially for small P (see Fig. 13 and 16). The difference arises from the fact that field configurations with a small perimeter mainly consist of single pixels in a large, homogeneously colored background. As a single pixel of different color contributes 4 perimeter sections to the overall perimeter, multiples of 4 are strongly favored for small perimeters. On the other hand, multiples of 2, which are no multiples of 4, are suppressed, as they can only be formed by connected regions. The binomial approximation does not value these geometric constraints. However, when computing the compatibility according to Eq. (25), one usually averages over several overestimated contributions (from perimeter lengths with an even value, which is no multiple of 4) as well as several underestimated contributions (from perimeter lengths with a value, which is a multiple of 4) and therefore Eq. (29) is still useful to quantify deviations. How well these two contributions average out and the overall goodness of this approximation is examined in Section 5.2.2, where the developed framework is tested on Monte Carlo simulations.

The complexity of $\Pr(\chi)$ is already indicated in Fig. 13. Elaborated studies did not show any approximation, which is helpful for the quantification of deviations. Therefore, the best way to constrain the probability that a deviation is actually a random fluctuation, is Chebyshev's inequality [10]:

$$\Pr(|\chi(\varrho) - \langle \chi \rangle(\varrho)| \geq k) \leq \frac{\mathrm{Var}[\chi](\varrho)}{k^2} \tag{30}$$

Using the expectation value from Eq. (23) and the variance from Eq. (26), this yields an upper limit on the probability of a certain fluctuation. Even though this constraint is rather loose, it can be used to reveal significant deviations from the background model.

Knowing the expected structure of Poisson fields and how to quantify deviations in the field structure, the developed techniques can be used to analyse measured data. Their application to H.E.S.S. sky maps is presented in the following sections.

[6] As periodic boundary conditions were imposed, there has to be a white region at the opposite side.

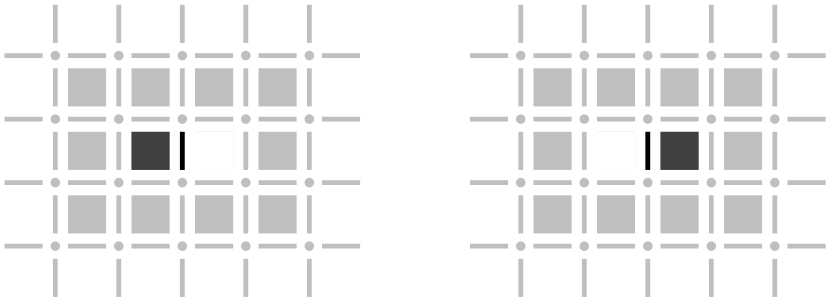

Fig. 14: *Possible pixel configurations and according probabilities to find an edge, which contributes to the perimeter. p denotes the probability of a pixel to be black.*

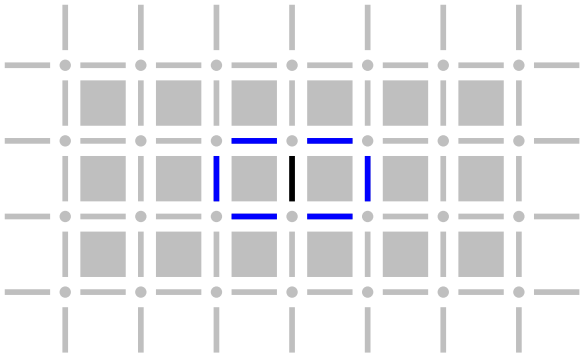

Fig. 15: *Correlation radius of an edge. The edges correlated to the black edge in the middle are highlighted.*

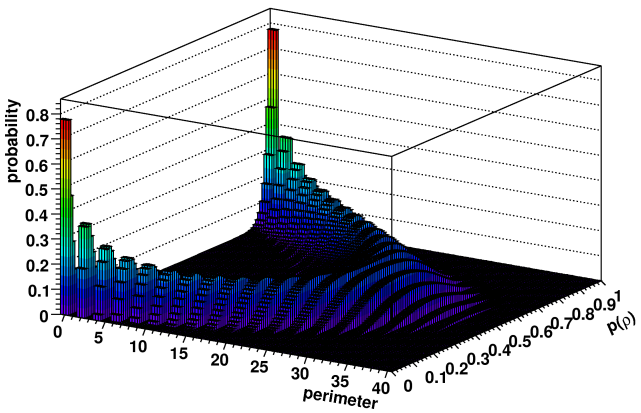

Fig. 16: *Binomial approximation to the distribution of the perimeter P for an image of 5×5 pixels as a function of the probability p of a pixel to be black.*

5 Minkowski Analysis of H.E.S.S. Data

While the standard H.E.S.S. analysis is optimised for the discovery of regions with a significant excess of photon-like events above the expected background rate, the new approach in this thesis is an analysis based on the structure of the data. Instead of looking for regions with unexpected high count rates, the morphology of the sky map is compared to the expected structure of pure background and sources are identified by significant deviations of the observed data from the expected behavior.

5.1 Acceptance Correction

The expected morphology of a sky map showing counts of gamma-like events differs between observations of different targets due to the camera acceptance of the experiment. Instead of having a homogeneous isotropic background one always ends up with photon count maps similar to Fig. 17. However, as the camera acceptance is a feature of the detector, there are ways to measure and model it (see Section 3.3). Based on an acceptance map like Fig. 18, one can apply an acceptance correction to the sky map, to regain an isotropic and homogeneous structure for background measurements. The naive approach to such an acceptance correction is to weigh each bin of the sky map with the corresponding inverse acceptance. Unfortunately, this leads to fractional photon counts and destroys the Poisson structure of the measurement. Therefore, the approach used here emanates from the hypothesis that the sky map already shows a Poisson field, but that the expected number of photons varies from bin to bin. Expecting to find a globally constant number of expected photon events λ for an ideal camera acceptance, the expected photon count λ_i for bin i has to be linked to the acceptance of the bin acc_i by the simple relation

$$\lambda_i = \mathrm{acc}_i \cdot \lambda. \tag{31}$$

Since the sum of two Poisson distributed random variables with the expectation values λ_1 and λ_2 is a Poisson distributed random variable with an expectation value of $\lambda_1 + \lambda_2$, a homogeneous isotropic Poisson field can be regained by adding to each bin a Poisson distributed randomly chosen number of photons with an expectation value of

$$\lambda_i^+ = (1 - \mathrm{acc}_i) \cdot \lambda. \tag{32}$$

In regions with high acceptance, one finds $\lambda_i^+ \approx 0$ and therefore the measured data stay effectively unchanged. The additionally created pseudo photon counts may cover features in regions of low acceptance, but do not introduce additional structural deviations from the homogeneous isotropic Poisson field. The fact that features of regions with low acceptance are covered beneath a substantial layer of Poisson distributed pseudo photon counts just accounts for the fact that the sensitivity of the telescope decreases with decreasing acceptance. Summarizing, the addition of random numbers to the measured data can actually help to analyze its structure.

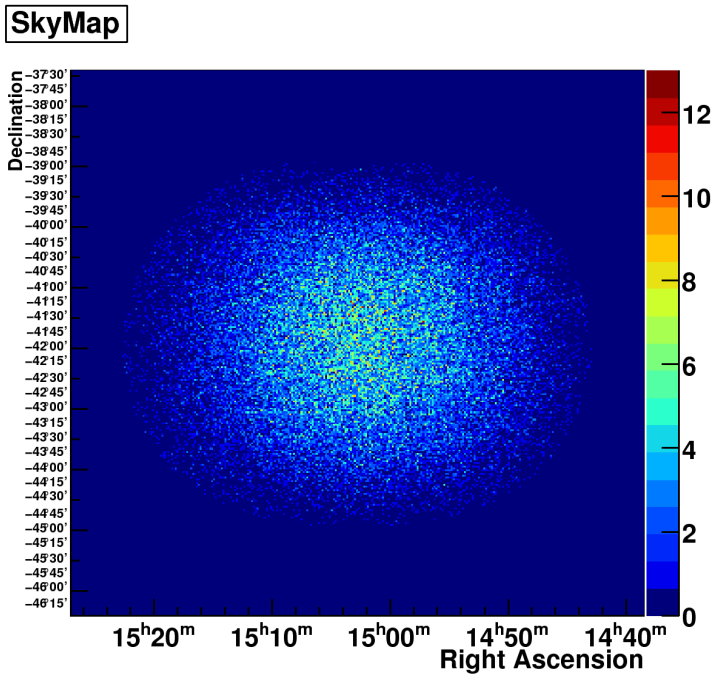

Fig. 17: *Typical H.E.S.S. sky map, showing in each bin the number of detected gamma-like events.*

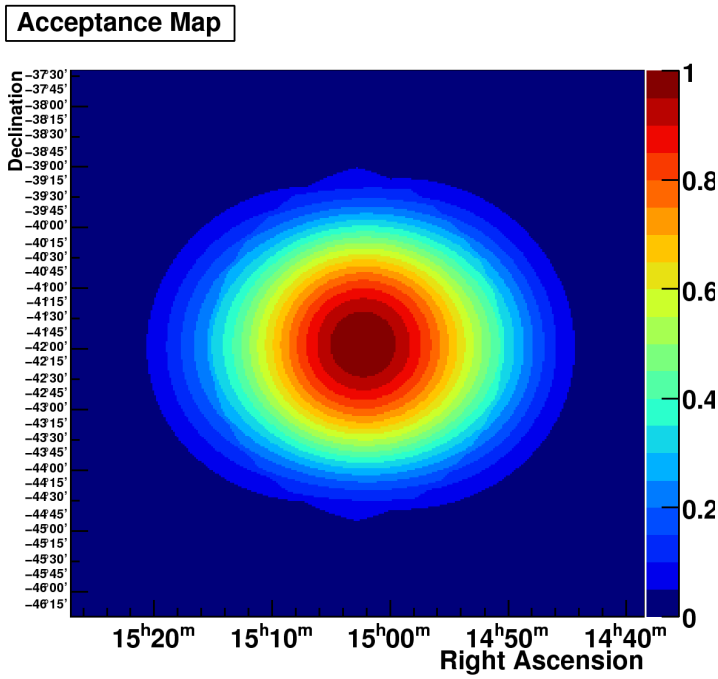

Fig. 18: *Typical H.E.S.S. acceptance map which shows the spacial sensitivity of the telescope for a given observation.*

5.2 Global Structure Analysis

A first application of the structure analysis using Minkowski functionals is the comparison of the global structure of a measured sky map to the structure of a Poisson field. To check the validity of the used background model, the first analysed H.E.S.S. data is a background data set. The background set comprises seven runs, which did not exhibit any signs of gamma-ray sources in careful studies using the H.E.S.S. standard analysis. The chosen runs correspond to a live time of about 3 h with a total number of 70918 gamma-like events. Instead of evaluating the whole sky map of these observations, the Minkowski functionals are computed for a region of 100×100 bins centered at the target position (red region in Fig. 19). In this inner part, the camera acceptance does not drop below 20% and thus, one avoids evaluating regions without any photon counts at the edge of the field of view. To be able to apply the acceptance correction to the selected area, the parameter λ (see Eq. (31)) has to be estimated. This is done by weighing the content of each bin in the inner region sky_i with the inverse of its corresponding acceptance acc_i and finally determining the mean bin content of this weighed sky map:

$$\lambda = \frac{\sum_i {}^{\mathrm{sky}_i}/_{\mathrm{acc}_i}}{\sum_i 1}, \tag{33}$$

where the sum runs over all bins in the selected inner region. In the case of sources within the field of view, the regions around the sources have to be excluded from the considered bins in order to provide an adequate estimate of the background structure.

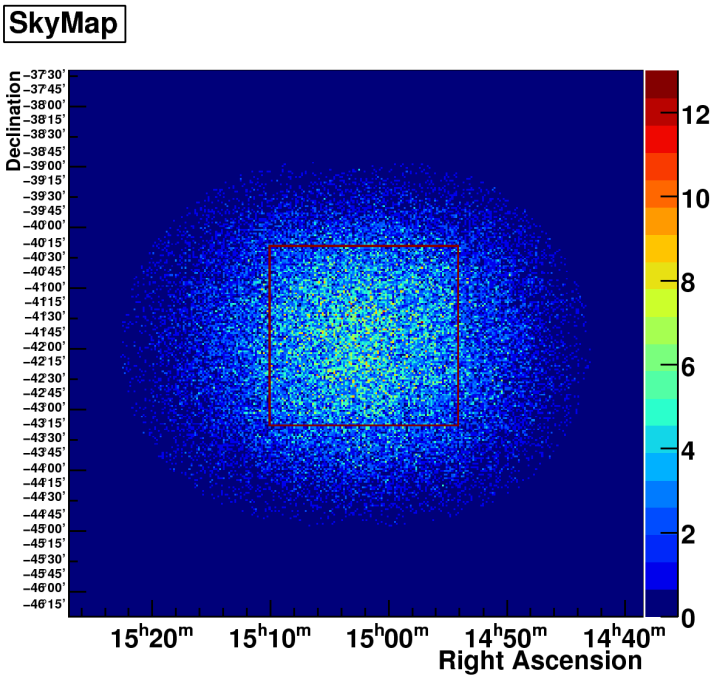

Fig. 19: *H.E.S.S. sky map of background data. The red rectangle marks the interior* 100×100 *bins region.*

Fig. 20 shows the change of the inner region due to the acceptance correction. The acceptance corrected sky map can then be used to determine the structure of the inner region. For the chosen background set the shape of the Minkowski functionals is in very good agreement with the theoretically predicted behavior (see Fig. 21). Using the methods introduced in Eq. (25), (29) and (30), the deviations in the structure of the whole inner region from an ideal Poisson background field can be quantified. As most of the fluctuations have an estimated probability of more than 10% to appear in a pure background measurement and none lie below 1% (see Fig. 22), the deviations can be well explained by statistical fluctuations. All in all this shows that the introduced methods succeed in describing background as measured by the H.E.S.S. experiment.

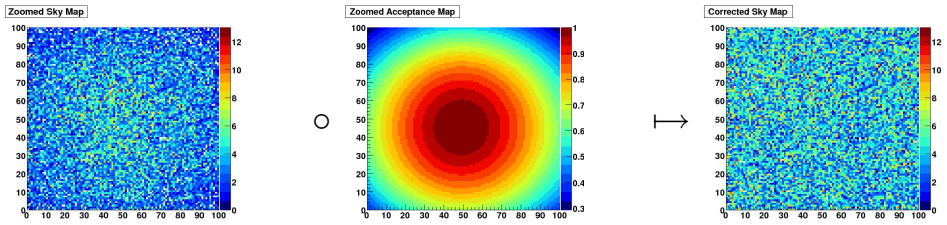

Fig. 20: *Acceptance correction for the inner* 100×100 *bins of the sky map from Fig. 19.*

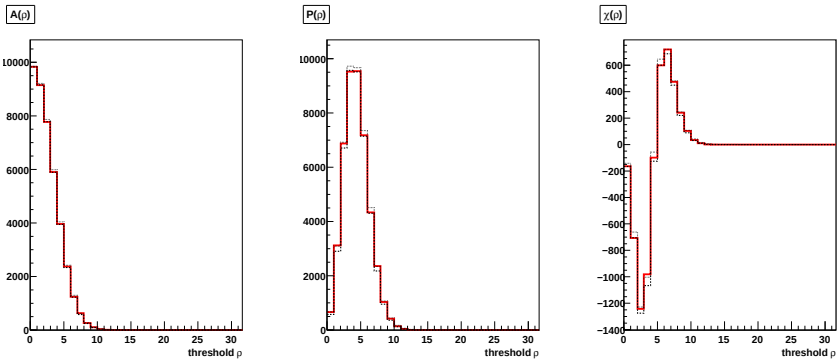

Fig. 21: *Minkowski functionals for background data (red) and* 1σ*-bands of theoretical prediction (dashed) according to Eq.* (23) *and* (26)

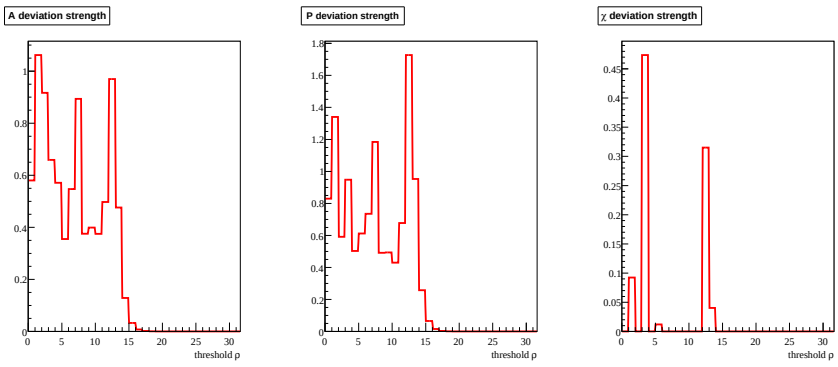

Fig. 22: *Negative decadic logarithm of the probability of the measured background structure to be compatible with a pure Poisson background.*

5.2.1 Tests on different Source Types

To see how sensitive the global structure analysis is to sources in the field of view, the methods described above are applied to data sets with known sources of different topologies.

First the influence of a very strong point source is studied. For this purpose, a flare of PKS 2155 − 304 is chosen. PKS 2155 − 304 is one of the best-studied blazers in the southern hemisphere and in July 2006 the H.E.S.S. collaboration witnessed an extraordinary outburst of this active galactic nucleus [13]. During 1.32 h of observation, a very strong signal with a significance level of about 150 σ has been observed. The global structure of the flare data is shown in Fig. 23 and indeed exhibits strong deviations from the expected curve shapes. As expected for a point source, the overall deviations in the structure are negligible except for the range of high thresholds, where the source is still visible, while background is expected to vanish.

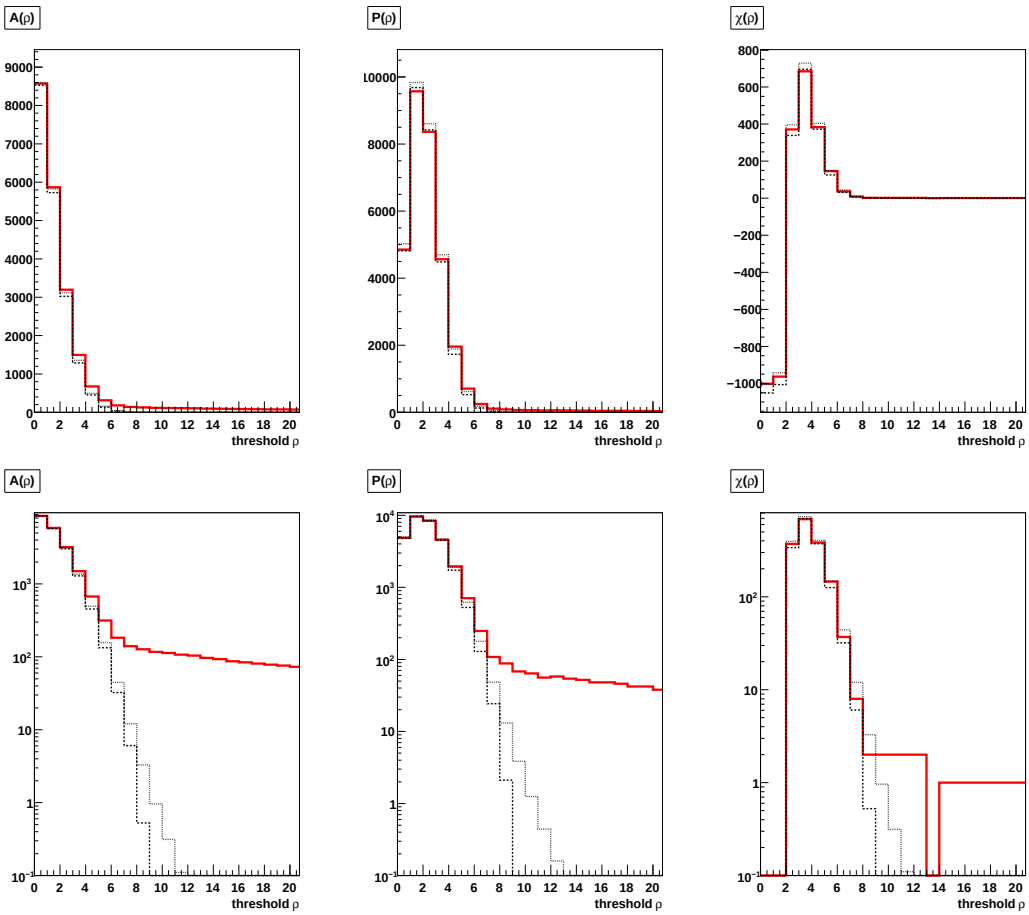

Fig. 23: *Minkowski functionals for* PKS 2155 − 304 *flare data (red) and* 1σ*-bands of theoretical prediction (dashed) according to Eq.* (23) *and* (26) *in linear (top) and logarithmic (bottom) scale.*

Computing the probability to find a statistical fluctuation at least as strong as the measured deviations in a pure Poisson background field one finds highly significant indications for an additional structure in the field of view. Defining the strength of a deviation to be the negative decadic logarithm of its probability to be compatible with a background fluctuation, one finds the deviation strength of PKS $2155-304$ for different thresholds ϱ distributed as shown in Fig. 24. The distributions of the deviation strength estimate the probability for the PKS $2155-304$ flare to be compatible with a background fluctuation to be less than 10^{-300}. The deviation strength of the area functional is only shown up to relatively small thresholds, as the compatibility probability for higher thresholds is so tiny that for numerical reasons it can no longer be calculated. As even Chebyshev's inequality indicates a significant deviation in the structure of the Euler characteristic, the global structure analysis is clearly sensitive to strong point sources.

Another object studied in this thesis is RX J1713.7 − 3946, a very extended supernova remnant with about twice the size of the full moon on the night sky [8]. The analyzed data consists of 170 runs with a live time of about 53 h. Computing the Minkowski functionals of RX J1713.7 − 3946 leads to the curves shown in Fig. 25. It is evident from Fig. 25 that an extended source is capable of changing the curve shapes significantly. The probability for this structure to be a random fluctuation is as low as 10^{-160}, proving the capability of the global structure analysis to identify extended structures using the Minkowski functionals (see Fig. 26).

The region of the Galactic Center is studied as last the example of source topologies. In addition to two distinct point sources, extended diffuse gamma-ray emission can be observed from this region [9]. The data set used for the global structure analysis consists of 129 runs with a live time of 51.5 h. The complex structure of the Galactic Center region is reflected in significant deviations of the Minkowski functionals from the background structure (see Fig. 27). The deviations are severe enough to show in the estimate for the Euler characteristic by Chebyshev's inequality (see Fig. 28). Thus, the developed techniques are especially suited to detect complex structures within the field of view.

The application of the Minkowski functionals to the H.E.S.S. sky maps showed that this technique is quite helpful to identify significant deviations from a pure Poisson background. However, in order to see whether the analysis succeeds not only qualitatively but quantitatively, one needs to test the used approximations with Monte Carlo simulations.

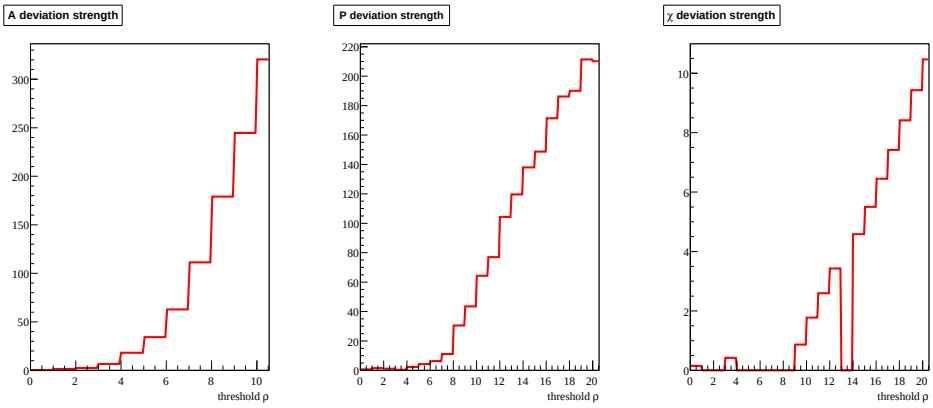

Fig. 24: *Deviation strength of* PKS 2155 − 304 *flare (Negative decadic logarithm of the probability of the measured* PKS 2155 − 304 *flare structure to be compatible with a pure Poisson background).*

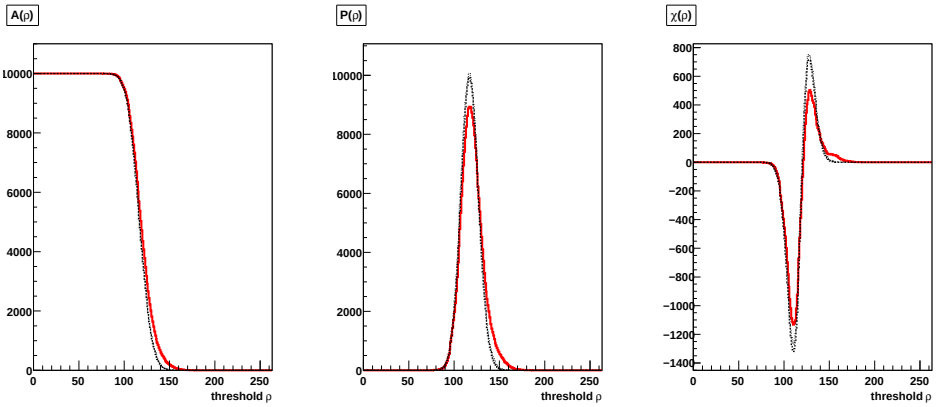

Fig. 25: *Minkowski functionals for* RX J1713.7 − 3946 *data (red) and* 1σ*-bands of theoretical prediction (dashed) according to Eq*(23) *and* (26).

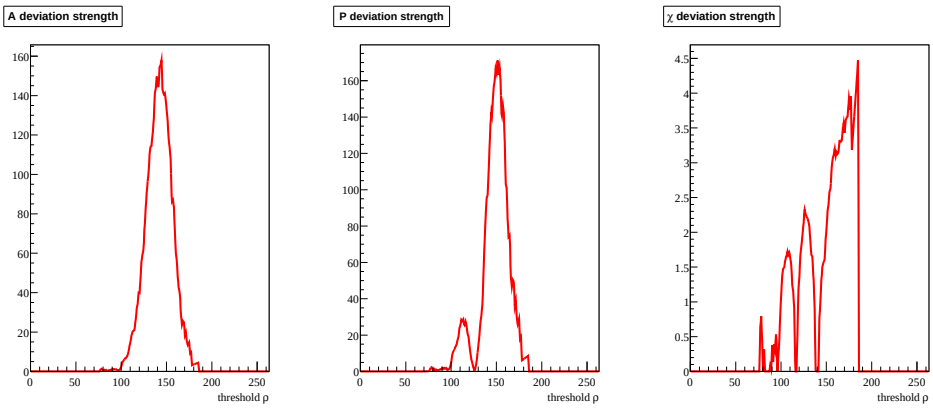

Fig. 26: *Deviation strength of* RX J1713.7 − 3946.

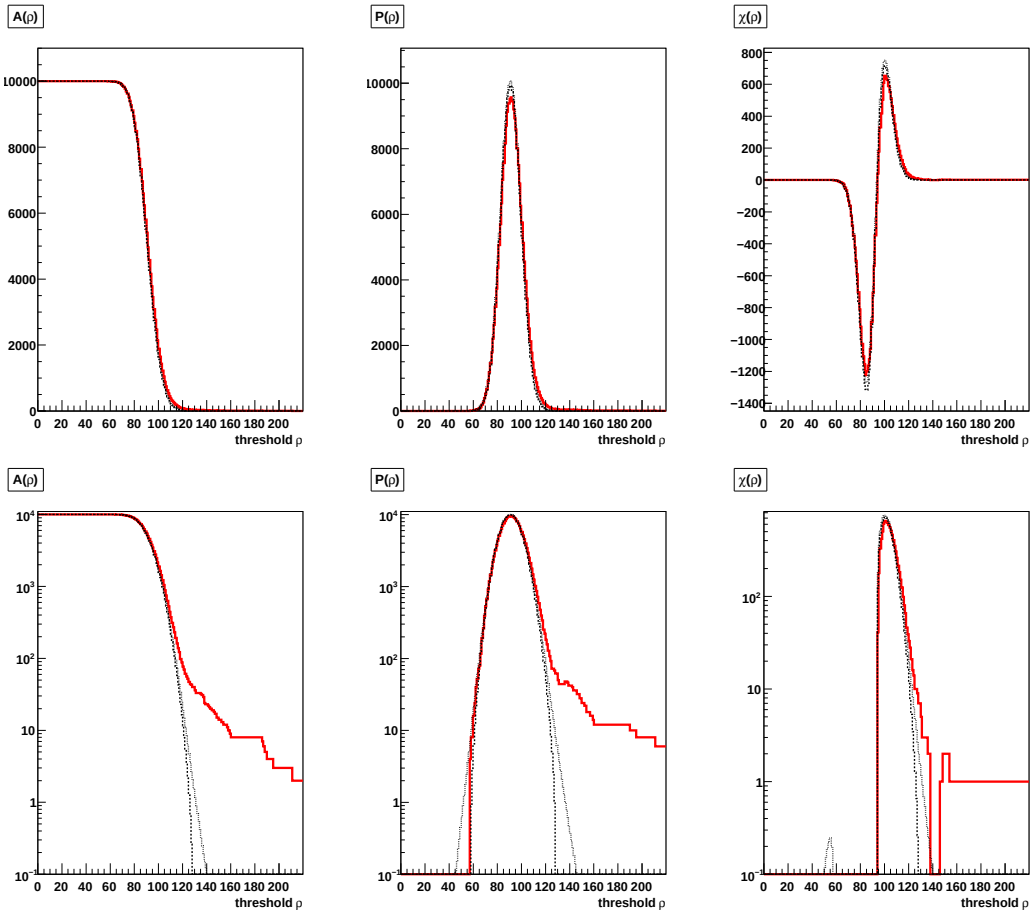

Fig. 27: *Minkowski functionals for data of the Galactic Center (red) and* 1σ*-bands of theoretical prediction (dashed) according to Eq.* (23) *and* (26) *in linear (top) and logarithmic (bottom) scale.*

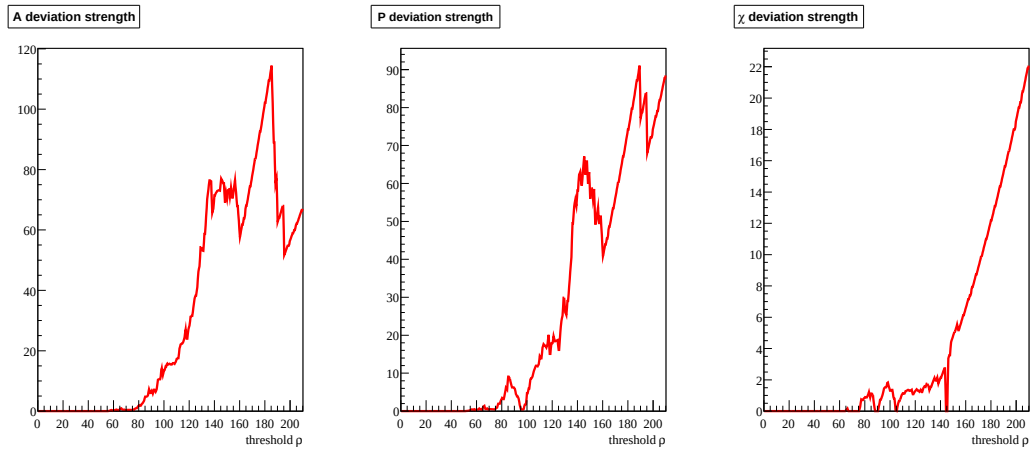

Fig. 28: *Deviation strength of the Galactic Center region.*

5.2.2 Tests on Monte Carlo Simulations

While the good agreement of the global structure analysis of H.E.S.S. background data with the underlying model and the sensitivity to the examined sources proves the general concept of the proposed technique, the question of the reliability needs further studies. In order to see, how robust the provided methods are regarding statistical fluctuations and how well the used approximations work, the global structure analysis is tested with Monte Carlo simulations. The simulations generate homogeneous and isotropic Poisson fields of 100×100 bins, filled with events from a pseudo-random number generator. Skipping the acceptance correction, these fields are analyzed using the same tools as used in the global structure analysis of the H.E.S.S. data.

A set of 500 simulated fields is processed for $\lambda = 4.14$ and $\lambda = 117$, which corresponds to the expected amount of photons in the analyzed data sets of the background measurement and RX J1713.7 − 3946 (see Section 5.2.1). The resulting distributions of the deviation strength are shown in Fig. 29. Since the exact distribution of the area functional is used to estimate the according deviation strength, the resulting frequency distributions show the expected behaviour. Independently of λ the majority of the fields is estimated to be compatible with the Poisson structure with a probability of at least 10%. The number of fields with higher deviation strengths drops nearly exponentially with a mean deviation strength of $\approx 0.6 \pm 0.4$, which varies only slightly over the range of thresholds.

The frequency distribution of the deviation strength regarding the perimeter functional reassembles the reference structure given by the area functional quite closely. Most fields are estimated to be Poisson fields with a probability of more than 10% and the number of fields with stronger deviations drops nearly exponentially. However, the mean strength of a deviation is estimated to be $\approx 0.9 \pm 0.6$ for very high or very low thresholds. Thus, the binomial approximation to the distribution of the perimeter functional (see Eq. (29)) tends to overestimate deviations in certain threshold ranges. In order to fully describe these systematic effects, further Monte Carlo studies with significantly higher statistics are needed.

As expected, the estimates for the Euler characteristic as provided by Chebyshev's inequality yield rather loose constraints on the compatibility with a Poisson structure. Virtually all of the examined Poisson fields are estimated to be featureless with probabilities above 10%. Since the usual deviation strength lies at $\approx 0.1 \pm 0.2$ for all thresholds, only very distinctive structures can be found using this approximation. Thus, the sensitivity to sources is very limited, but signals found using the global structure analysis with the Euler characteristic show a high confidence level.

All in all the global structure analysis works very well, but care has to be taken interpreting the deviation strength derived from approximate distributions of the Minkowski functionals. Further extensive studies with Monte Carlo simulations may help to correct the systematic errors in the results of the perimeter and improve the sensitivity of the Euler characteristic.

The global structure analysis provides neither the exact position of a source in the field of view nor information about its shape. Still it is very efficient in indicating additional structures in a sky map. As the computational effort for the global structure analysis is not very high, this method can be used to quickly check a data set for possible sources and help to decide if a detailed follow up analysis should be performed.

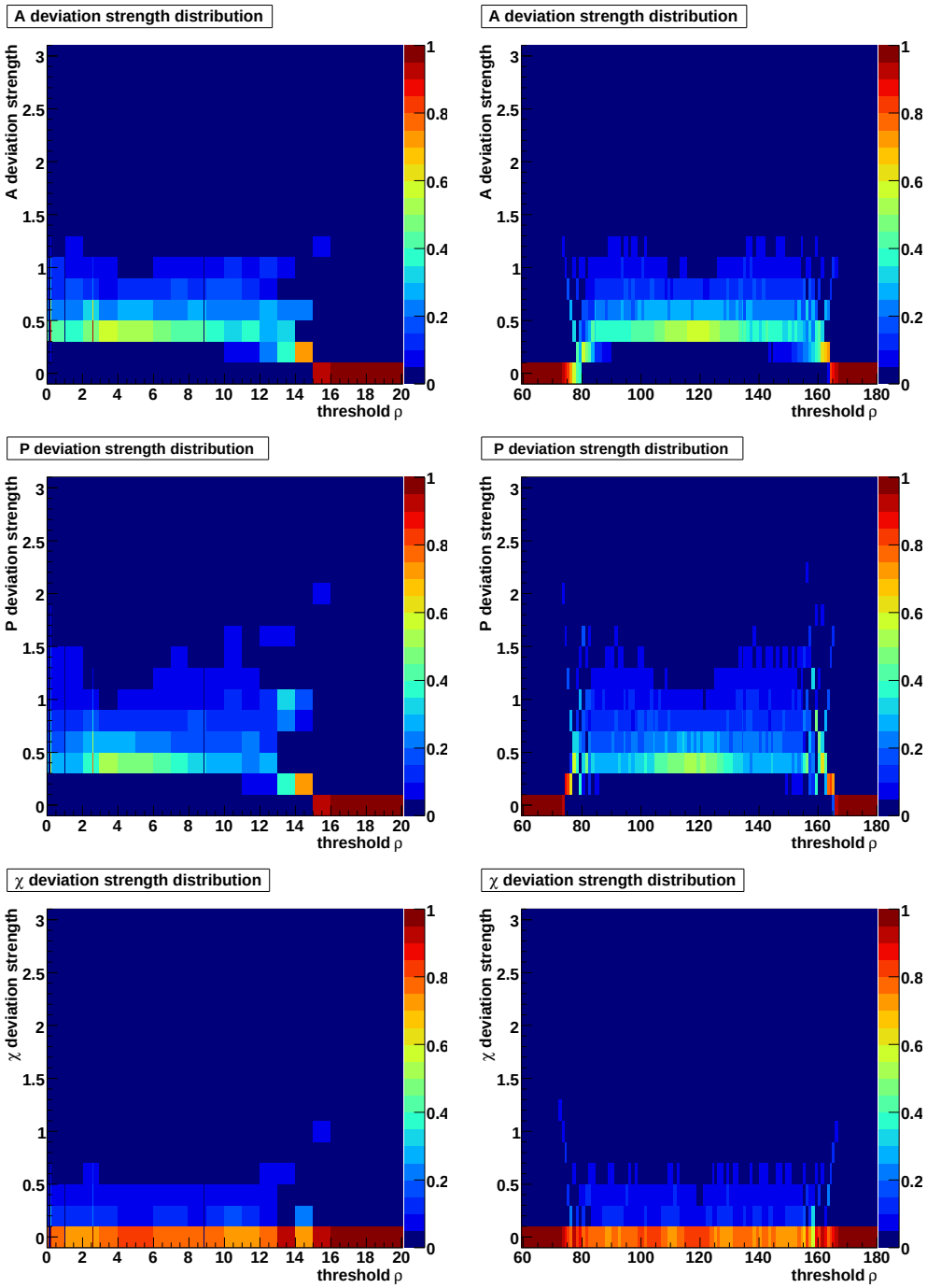

Fig. 29: *Relative frequencies of the deviation strength of the simulated background fields. Each slice for a certain value of ϱ gives the frequency distribution of the deviation strength of the* 500 *Poisson fields for the selected threshold. On the left the distributions for* $\lambda = 4.14$ *are shown and on the right the corresponding distributions for* $\lambda = 117$.

5.3 Minkowski Sky Maps

As the Minkowski functionals are motion-invariant, the exact location of a source in a sky map can not be derived from the global structure analysis. Instead, one can slide a small search window across the sky map and thus determine regions of exceptional structure by analyzing the content of the search window. This window should not be chosen too small, to give the interior some room to exhibit a structure at all. As the 5×5 field is the biggest field for which all distributions are easily accessible and exactly known, a scan window of this size is used for the further investigations. The main idea is to create a Minkowski sky map by centering the scan window on each bin of the selected inner region of the field of view and computing the local structure of the 5×5 field around the bin. The deviation strength of each local window is determined and used to fill the center bin with an according value. This way a sky map is created, which visualizes for each bin the structural deviations of its immediate vicinity from the expected behavior of a background field.

The main procedure is demonstrated using the example of RX J1713.7 − 3946 (see Section 5.2.1). The model parameter λ is estimated as usual from the entire inner region, excluding the source itself and the acceptance correction is applied to the studied region. Thereafter, the exact distributions of the functionals in the 5×5 field are used to quantify possible deviations. This leads to threshold dependent deviation strength similar to Fig. 22. In order to map these curves to a single value for each functional, which then can be filled into the corresponding Minkowski sky map, one can use two different approaches:

- calculating the mean of the curve by integration
- choosing the global maximum of the curve

Fig. 30 opposes the results of these two approaches for the evaluation of the area functional of RX J1713.7 − 3946.

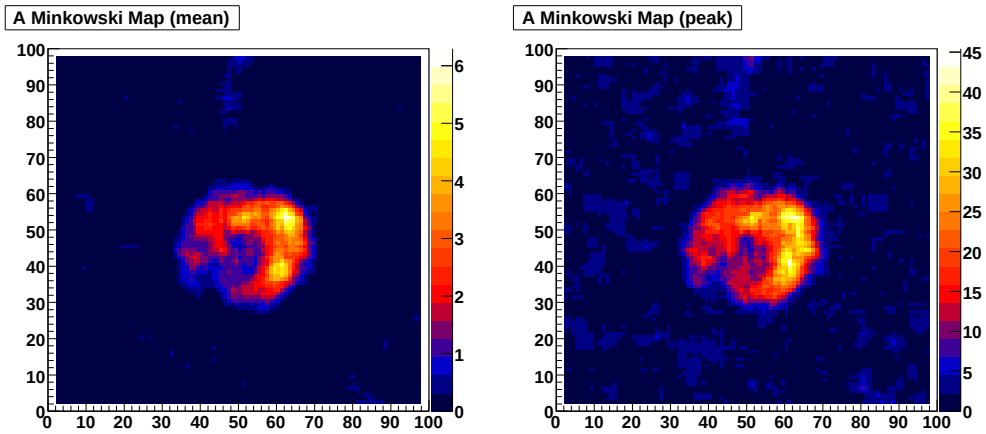

Fig. 30: *Minkowski sky maps for the area functional of* RX J1713.7 − 3946 *showing the mean (left) and the peak value (right) of the deviation strength distribution.*

While the mean value of the deviation strength distribution is less sensitive to fluctuations, it suffers from two disadvantages: A pronounced structure influences the functionals over a large range of threshold values ϱ. Thus the deviations in this range are highly correlated and a simple average may fail to quantify the overall deviation. Additionally it is not immediately

clear how to interpret the computed mean, as its value depends strongly on the threshold range used for integration. The main problem in this context is to restrict the considered threshold range. Without this restriction one could scale down the mean value to arbitrarily small values by averaging over regions where no deviations are expected. However, as the range of thresholds where deviations can be expected varies from data set to data set, it is difficult to define a universally valid integration region and thus, to obtain a properly normalized mean value.

The drawbacks of the peak value are the higher sensitivity to random fluctuations and the unknown position of the peak value. By choosing the maximum for each scan window, the threshold used to evaluate the window may be different for each bin. As Fig. 30 shows that both approaches lead to comparable results, the peak value is chosen for the further course of this work, as it is easier to interpret its value.

5.3.1 Comparison with standard H.E.S.S. Sky Maps

Fig. 31 shows the result of using the described methods to create Minkowski sky maps for the perimeter and the Euler characteristic of RX J1713.7 − 3946. The supernova remnant is most significantly detected in the Minkowski sky map of the area functional (see Fig. 30). The Minkowski sky maps of the perimeter and the Euler characteristic exhibit the same source morphology with less but still very high significance. A comparison of the found source morphology to the published H.E.S.S. analysis of RX J1713.7 − 3946 shows good agreement between the Minkowski sky maps and the H.E.S.S. standard analysis (see Fig. 32).

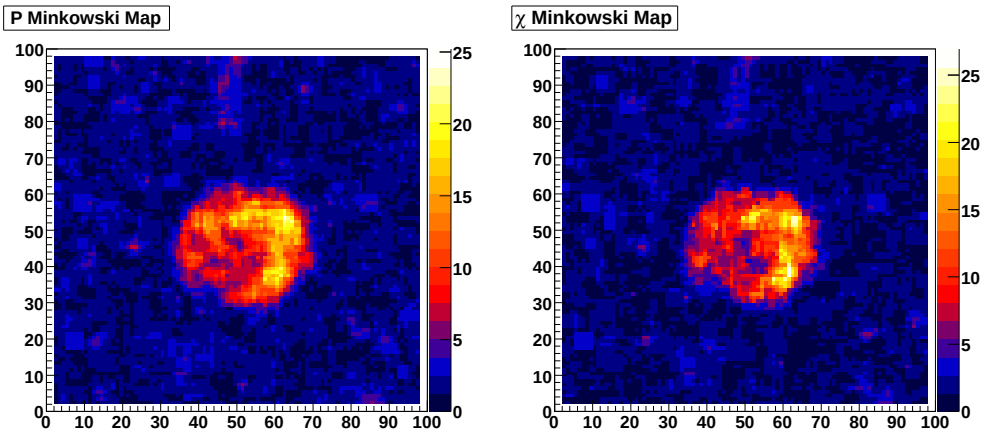

Fig. 31: *Minkowski sky maps of the perimeter functional (left) and the Euler characteristic functional (right) of* RX J1713.7 − 3946.

The complex structure of the Galactic Center makes this region another interesting candidate to be investigated with Minkowski sky maps (see Section 5.2.1). The results of the local structure analysis are shown in Fig. 33 and mainly reveal the location of the strong point source HESS J1745 − 290. Again this source is most significantly seen in the Minkowski sky map of the area functional, but also clearly visible in the maps of the perimeter and Euler characteristic. Since the color scale of the Minkowski sky maps in Fig. 33 is completely

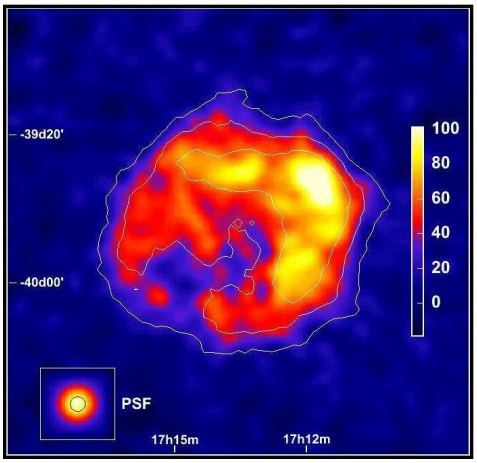

Fig. 32: *Excess map of* RX J1713.7 − 3946 *resulting from H.E.S.S. standard analysis [8]. The white contour lines indicate the significance of the observed signal and correspond to* $5\,\sigma$*,* $10\,\sigma$ *and* $15\,\sigma$*.*

dominated by the strong signal from HESS J1745 − 290, the scale has to be adjusted in order to see the second point source G0.9 + 0.1 and the band of diffuse emission from that region. Fig. 34 shows the same Minkowski sky maps with the color scale saturated at a value of 7, corresponding to deviations, which are less probable than a $5\,\sigma$ fluctuation in a Gaussian distribution. Here, the second point source is clearly visible and the extended structure due to the diffuse emission emerges. Thus all the structures found in that region by the standard H.E.S.S. analysis (see Fig. 35) can be identified using Minkowski sky maps.

In order to see how reliable structural deviations are quantified and if the Minkowski sky maps allow not only qualitative but quantitative studies of H.E.S.S. data, the methods are tested by Monte Carlo simulations.

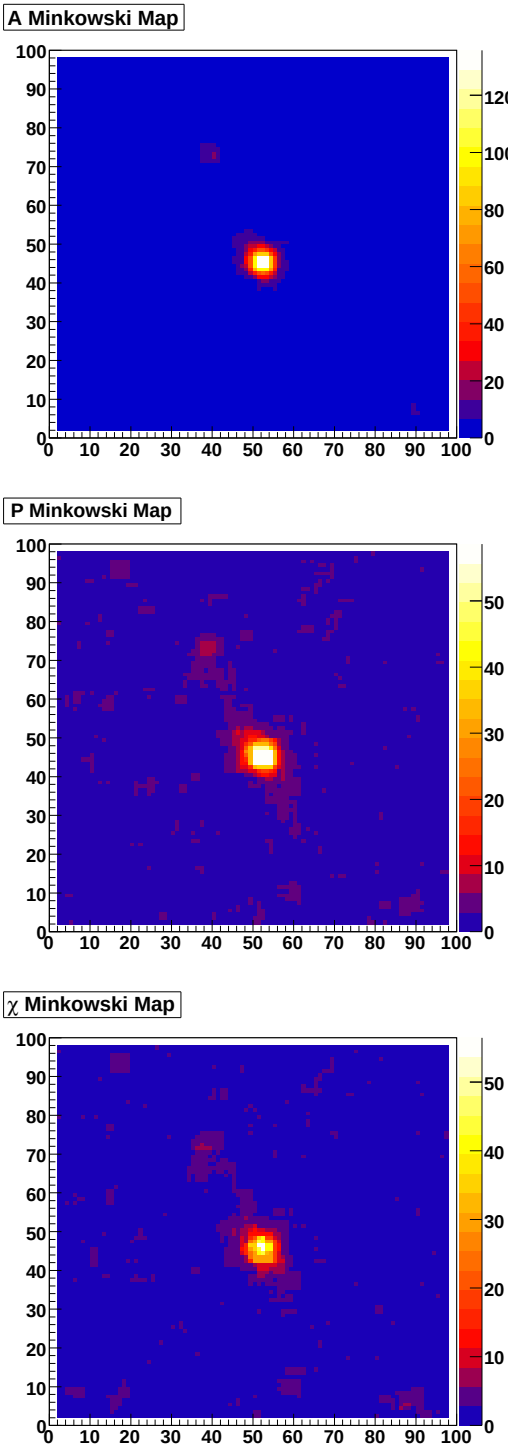

Fig. 33: *Minkowski sky maps of the Galactic Center region.*

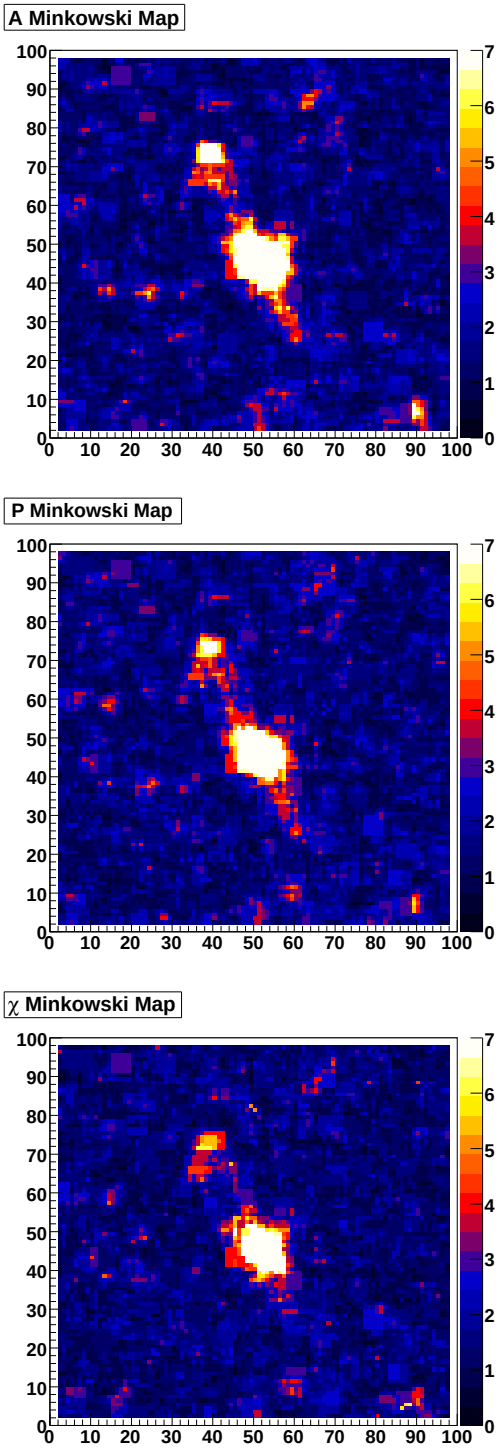

Fig. 34: *Minkowski sky maps of the Galactic Center region with the color scale saturated at a value of 7.*

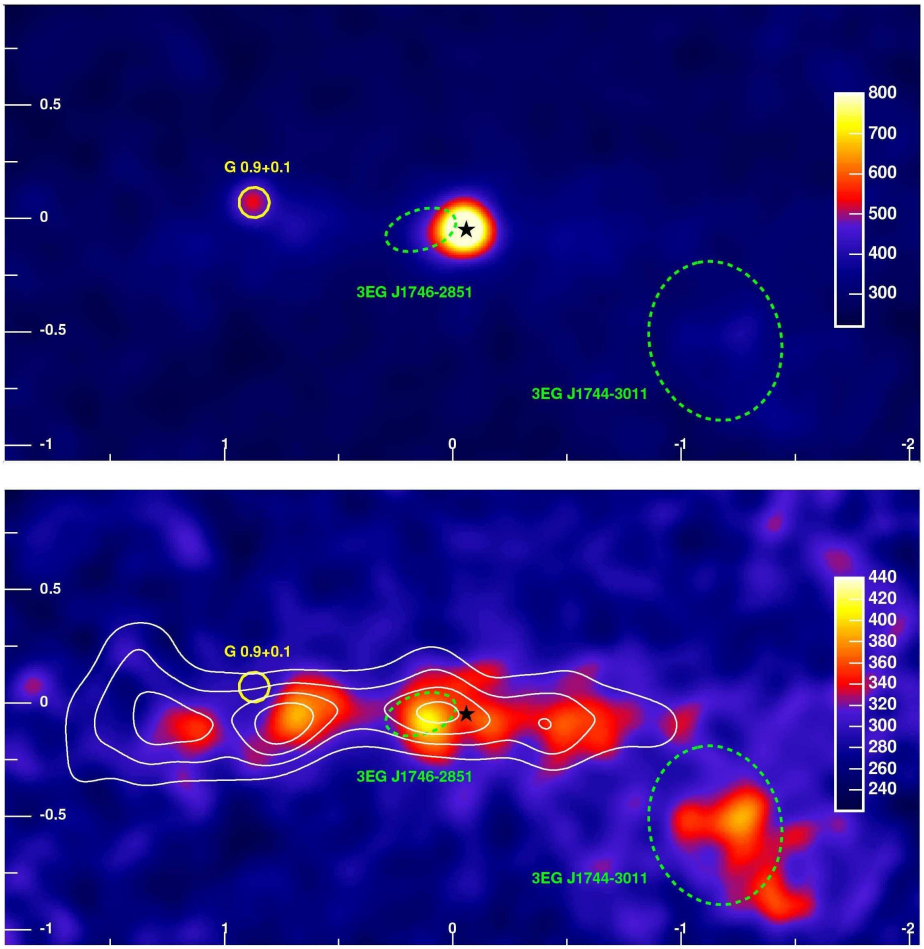

Fig. 35: *Excess maps of the Galactic Center region resulting from H.E.S.S. standard analysis [9]. The excess map on the top shows the two point sources* $\mathrm{G}0.9+0.1$ *and* $\mathrm{HESS\ J}1745-290$. *The bottom image shows the same region after substraction of the signals from the point sources.*

5.3.2 Monte Carlo Studies of Minkowski Sky Maps

To verify the the methods used to analyze the local structure, Minkowski sky maps of the Poisson background fields simulated in Section 5.2.2 are computed as well using the mean value as using the peak value of the deviation strength distribution (see Fig. 30). For each simulated Poisson field the resulting value of each bin of the Minkowski sky map is filled in a histogram. Fig. 36 shows the resulting relative frequencies using the mean value of the deviation strength distribution. In addition to the arbitrary normalisation the shown distributions reveal further disadvantages in the use of the mean value of the deviation strength. The shapes of the curves differ for different parameters λ and the distributions flattens towards high deviation strengths. Both features pose severe problems for a reliable quantitative analysis of sky maps. In order to account for the different shapes, the right distribution would have to be obtained by separate Monte Carlo simulations for every needed value of λ, which leads to an enormous computational effort. The flattening of the distribution expresses the tendency of this approach to overestimate the strength of a deviation, leading to high systematic uncertainties. Summarizing, Minkowski sky maps showing the mean value of the local compatibility to a Poisson field are inappropriate for a quantitative analysis of local structures.

The distributions for the Minkowski sky maps showing the peak value of the deviation strength of the local structure are shown in Fig. 37. In contrast to Fig. 36, all six distributions of the peak values exhibit the same shape, independently of the parameter λ or the used Minkowski functional. Furthermore, there is no need to infer this global distribution from extensive Monte Carlo simulations, as all of the histograms agree very well with the indicated functional dependence of 10^{-x}. This facilitates an easy interpretation of the values in Minkowski sky maps: Given a bin i with the value mink_i in a Minkowski sky map, the probability to find the observed local structure in a Poisson field is given by 10^{-mink_i}.

As the shown results are based on a small sample of 500 Poisson fields, further extensive studies are needed to verify their validity. Nevertheless, the properties of Minkowski sky maps, which show the peak value of the local deviation strength, are very promising for an accurate evaluation of local structures.

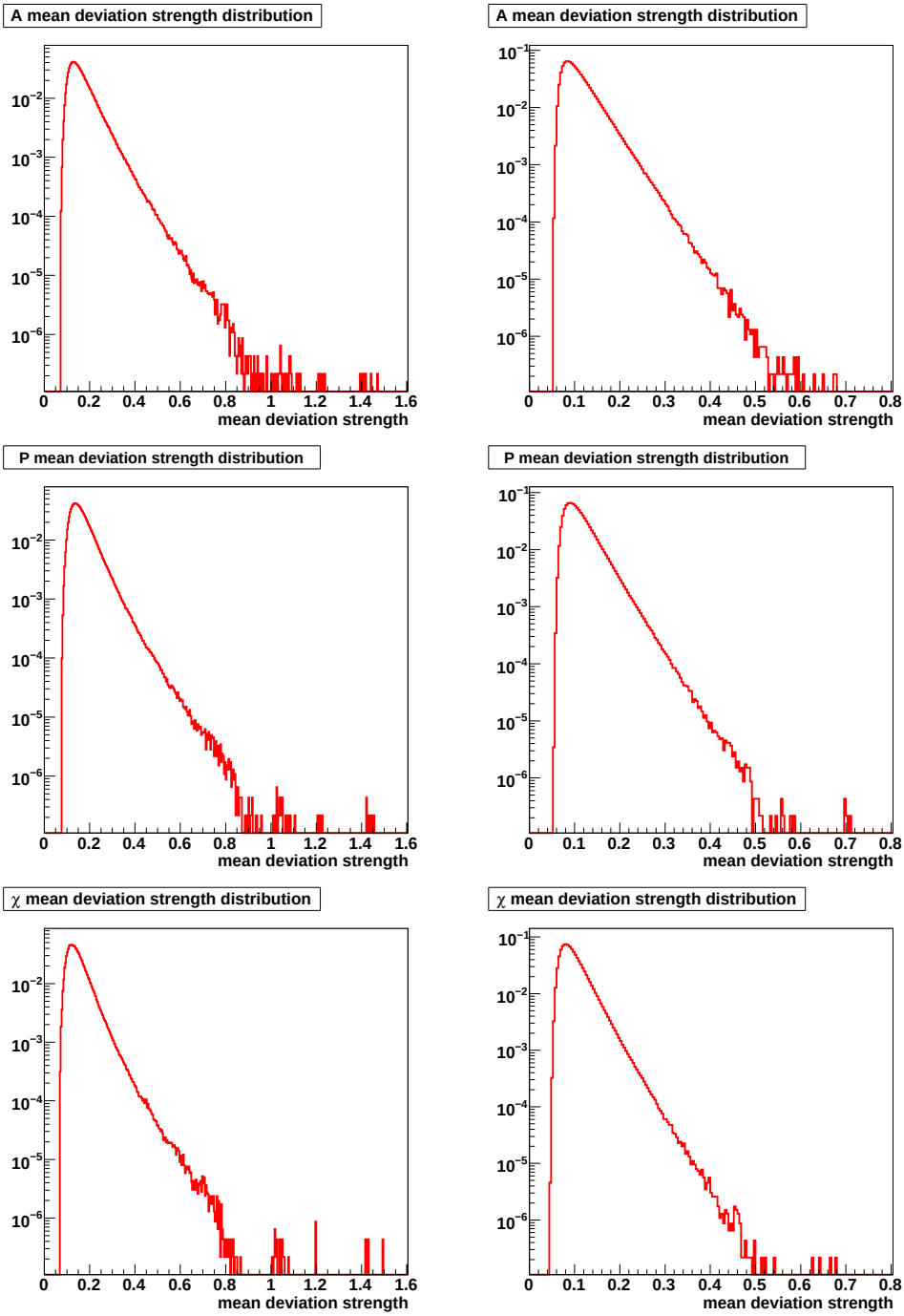

Fig. 36: *Relative frequency of the local deviation strength averaged over an arbitrary range of thresholds. On the left the distributions for $\lambda = 4.14$ are shown and on the right the corresponding distributions for $\lambda = 117$.*

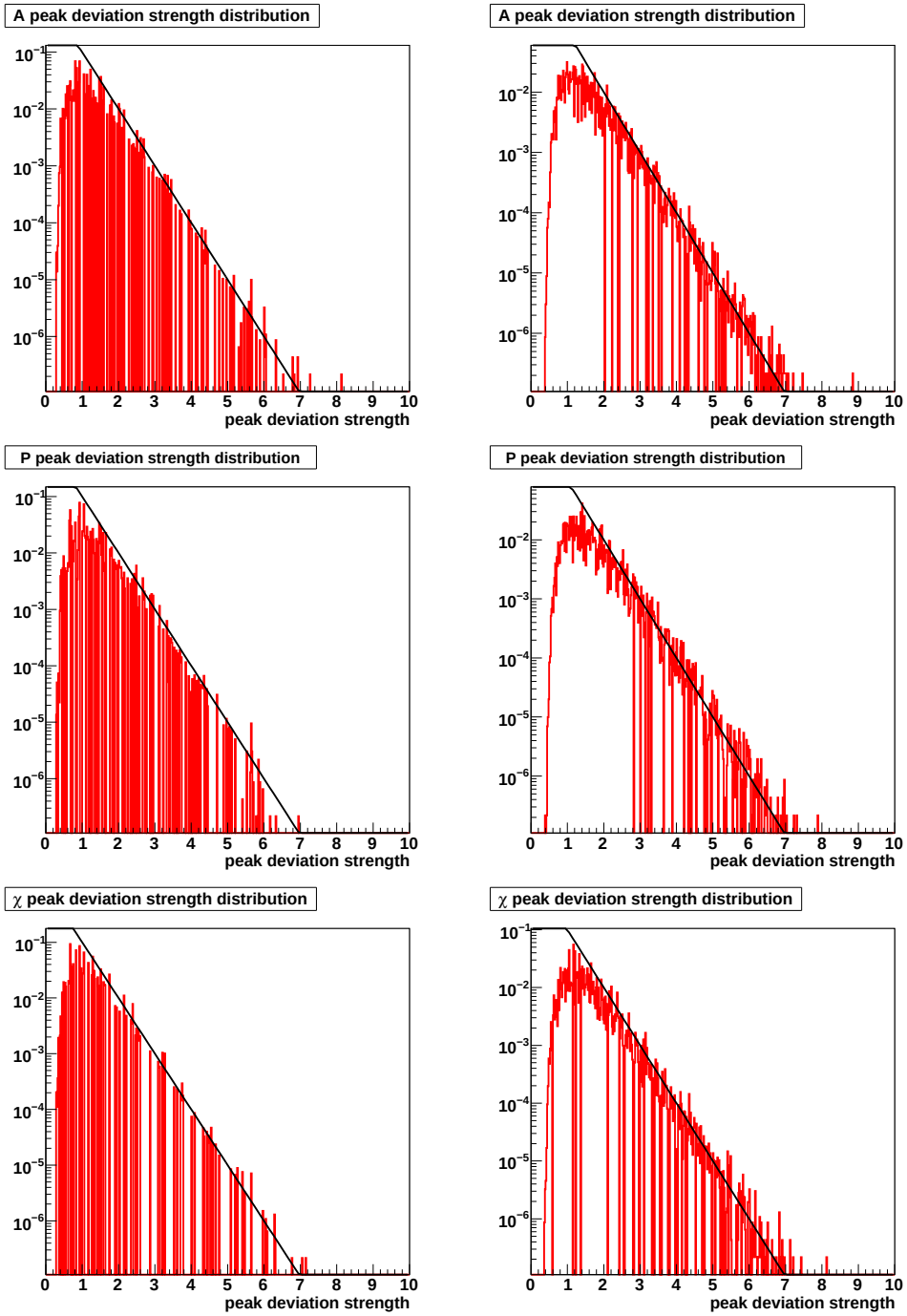

Fig. 37: *Relative frequency of the peak value of the local deviation strength. On the left the distributions for* $\lambda = 4.14$ *are shown and on the right the corresponding distributions for* $\lambda = 117$*. The black lines indicate the function* 10^{-x}*.*

6 Conclusion and Outlook

The background as detected by H.E.S.S. can be described by a homogeneous and isotropic Poisson field. Such fields exhibit a characteristic structure, which can be quantified using the Minkowski functionals. Sources on top of the background can be identified by deviations between the structure of the observed field and the structure of the expected Poisson field. This approach was tested on different source types, showing that the Minkowski functionals are sensitive to point sources as well as to extended structures and especially to complex configurations like multiple sources or diffuse emission. The exact position of a source and its morphology is accessible via Minkowski sky maps, which illustrate for each bin the compatibility of the local structure with a background field. All in all, the morphologies found by the Minkowski analysis are in good agreement with those found by the H.E.S.S. standard analysis. However, this first implementation of an analysis based on the structure of the observed data as described by the Minkowski functionals impressively demonstrates the feasibility of such an analysis.

The H.E.S.S. standard analysis focuses on exceptionally intense regions of gamma-like events in the observed sky. This corresponds merely to the regime of high thresholds of the area functional. Using the additional information of all the Minkowski functionals over the total range of thresholds promises a substantial gain in sensitivity to gamma-ray sources on top of the observed background.

Appendices

A Variances of Image Components

Be $a_i \in \{0, 1\}$ the color of the i^{th} pixel (with 0 labeling white pixels and 1 labeling black pixels). Then

$$\langle a_i \rangle = p, \tag{34}$$

with p being the probability of a pixel to be black (see Eq. (19)). Since

$$n_2 = \sum_i a_i, \tag{35}$$

one finds

$$\begin{aligned} \langle n_2 \rangle &= \sum_i \langle a_i \rangle \\ &= N^2 p, \end{aligned} \tag{36}$$

consistent with Eq. (22). In order to calculate the variance

$$\mathrm{Var}[n_2] = \langle n_2^2 \rangle - \langle n_2 \rangle^2, \tag{37}$$

$\langle n_2^2 \rangle$ has to be determined.

$$\begin{aligned} \langle n_2^2 \rangle &= \sum_{i,j} \langle a_i a_j \rangle \\ &= \sum_i \langle a_i^2 \rangle + \sum_{\substack{i,j \\ \text{correlated}}} \langle a_i a_j \rangle + \sum_{\substack{i,j \\ \text{independent}}} \langle a_i \rangle \langle a_j \rangle \\ &= N^2 p + \sum_{\substack{i,j \\ \text{correlated}}} \Pr(a_i \wedge a_j) + \sum_{\substack{i,j \\ \text{independent}}} p^2 \end{aligned} \tag{38}$$

With all pixels being independent, one finds

$$\begin{aligned} \langle n_2^2 \rangle &= N^2 p + 0 + (N^4 - N^2) p^2 \\ &= N^4 p^2 + N^2 p (1 - p). \end{aligned} \tag{39}$$

Consequently

$$\mathrm{Var}[n_2] = N^2 p (1 - p), \tag{40}$$

consistent with the binomial distribution of Eq. (24).

The derivation of $\mathrm{Var}[n_1]$ follows the same scheme. Be $e_i \in \{0, 1\}$ the color of the i^{th} edge (with 0 labeling white edges and 1 labeling black edges). Using Eq. (20), one finds

$$\langle e_i \rangle = p(2 - p). \tag{41}$$

Similar to Eq. (38), it holds

$$\langle n_1^2 \rangle = 2N^2 p(2-p) + \sum_{\substack{i,j \\ \text{correlated}}} \Pr(e_i \wedge e_j) + \sum_{\substack{i,j \\ \text{independent}}} p^2(2-p)^2. \tag{42}$$

As shown in Fig. 15, for each i there are 6 corresponding js, which are correlated. Fig. 38 shows that the probability for two correlated edges to be black is given by

$$\Pr(e_i \wedge e_j) = p(1+p-p^2). \tag{43}$$

Therefore,

$$\langle n_1^2 \rangle = 2N^2 p(2-p) + 2N^2 \cdot 6 \cdot p(1+p-p^2) + (4N^4 - 14N^2)p^2(2-p)^2, \tag{44}$$

and thus,

$$\mathrm{Var}[n_1] = 2N^2 p(1-p)^2(8-7p), \tag{45}$$

which is in perfect agreement with the formula found for $N = 5$ (see Eq. (27)).

Finally, be $v_i \in \{0, 1\}$ the color of the i^{th} vertex (with 0 labeling white vertices and 1 labeling black vertices). Using Eq. (21), one finds

$$\langle v_i \rangle = p(2-p)(2-2p+p^2). \tag{46}$$

The correlated vertices can be divided into two classes, as shown in Fig. 39. Each vertex is correlated with 4 vertices from each class and the probability to find 2 correlated black vertices is given by

$$\Pr(v_i \wedge v_j) = \begin{cases} p(1+9p-27p^2+33p^3-21p^4+7p^5-p^6) & \text{for class 1} \\ p(2-p)(1+2p-5p^2+4p^3-p^4) & \text{for class 2} \end{cases} \tag{47}$$

Thus,

$$\begin{aligned} \langle n_0^2 \rangle = & N^2 p(2-p)(2-2p+p^2) + \\ & N^2 \cdot 4 \cdot p(1+9p-27p^2+33p^3-21p^4+7p^5-p^6) + \\ & N^2 \cdot 4 \cdot p(2-p)(1+2p-5p^2+4p^3-p^4) + \\ & (N^4 - 9N^2)p^2(2-p)^2(2-2p+p^2)^2, \end{aligned} \tag{48}$$

and finally

$$\mathrm{Var}[n_0] = N^2 p(1-p)^4(16-38p+32p^2-9p^3). \tag{49}$$

Therefore, the formulas found for small fields given in Eq. (27) are valid for arbitrary field sizes.

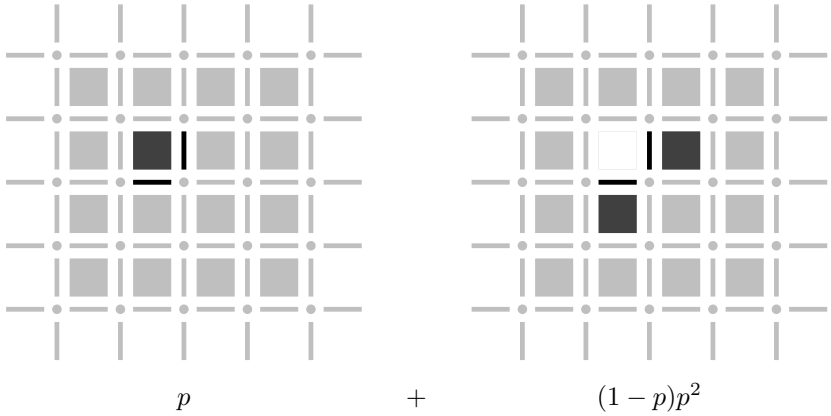

Fig. 38: *Probability of two correlated edges to be black. p denotes the probability to find a black pixel.*

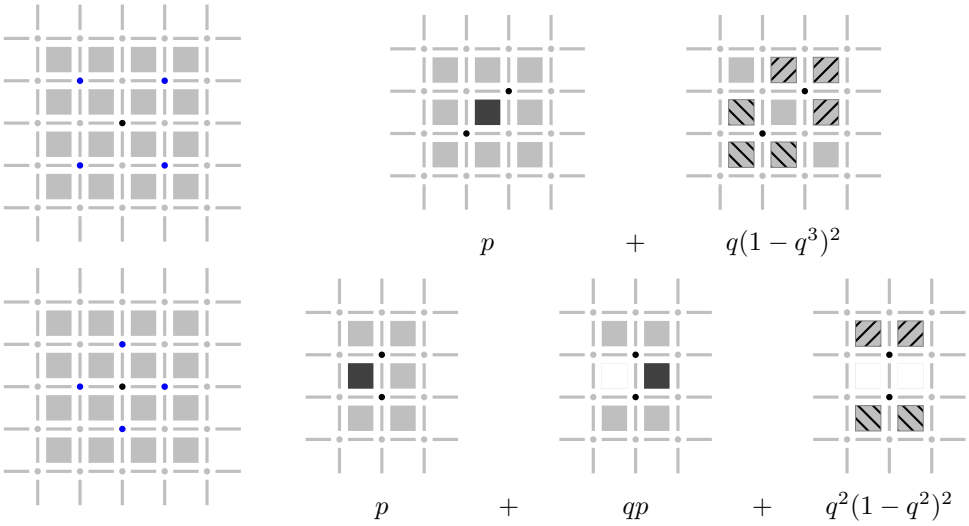

Fig. 39: *The two classes of correlated vertices (top and bottom) and their associated probability of two correlated vertices to be black. p denotes the probability of a black pixel to be black and $q = 1 - p$ denotes the probability of a pixel to be white.*

References

[1] F Aharonian, J Buckley, T Kifune, and G Sinnis. High energy astrophysics with ground-based gamma ray detectors. *Reports on Progress in Physics*, 71(9):096901 (56pp), 2008.

[2] A. M. Hillas. Cerenkov light images of EAS produced by primary gamma. In F. C. Jones, editor, *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, pages 445–448, August 1985.

[3] D. Berge. *The gamma-ray supernova remnant RX J1713.7-3946 with H.E.S.S.* PhD thesis, Ruperto-Carola University of Heidelberg, February 2006.

[4] W. Hofmann, I. Jung, A. Konopelko, H. Krawczynski, H. Lampeitl, and G. Puehlhofer. Comparison of techniques to reconstruct VHE gamma-ray showers from multiple stereoscopic Cherenkov images. *Astroparticle Physics*, 12:135–143(9), November 1999.

[5] H. E. S. S. Collaboration. Observations of the Crab nebula with HESS. *Astronomy and Astrophysics*, 457(3):899–915, October 2006.

[6] D. Berge, S. Funk, and J. Hinton. Background modelling in very-high-energy gamma-ray astronomy. *Astronomy and Astrophysics*, 466(3):1219–1229, May 2007.

[7] T.-P. Li and Y.-Q. Ma. Analysis methods for results in gamma-ray astronomy. ApJ, 272:317–324, September 1983.

[8] H. E. S. S. Collaboration. A detailed spectral and morphological study of the gamma-ray supernova remnant RX J1713.7-3946 with HESS. *Astronomy and Astrophysics*, 449(1):223–242, April 2006.

[9] H. E. S. S. Collaboration. Discovery of Very-High-Energy Gamma-Rays from the Galactic Centre Ridge. *Nature*, 439:695–698, February 2006.

[10] Ilja N. Bronstein, Konstantin A. Semendjajew, and Gerhard Musiol. *Taschenbuch der Mathematik*. Verlag Harri Deusch, fourth edition, 1999.

[11] H. Hadwiger. *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*. Springer Verlag, 1957.

[12] K. Michielsen, H. De Raedt, and J.Th.M. De Hosson. Aspects of Mathematical Morphology. *Advances in Imaging and Electron Physics*, 125:119 – 195, 2002.

[13] H. E. S. S. Collaboration. An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304. *The Astrophysical Journal Letters*, 664(2):L71–L74, 2007.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich bei der Anfertigung dieser Diplomarbeit unterstützt haben. Insbesondere bei

- meinen Eltern, für die enorme Geduld und bedingungslosen Rückhalt.
- Prof. Christian Stegmann, für die inspirierenden Diskussionen zu später Stunde, die gewährten Freiheiten in der Bearbeitung des Themas und die umfassende Unterstützung.
- Prof. Klaus Mecke, für die Hilfe bei der Einarbeitung in das Thema und den wertvollen Meinungsaustausch im weiteren Verlauf der Arbeit.
- Julia Brucker, für zahlreiche nützliche Ratschläge und ihre große Hilfsbereitschaft.
- Fabian Schöck, für viele hilfreiche Anmerkungen zu dieser Arbeit.

Erklärung

Hiermit bestätige ich, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Erlangen, den 12. Dezember 2008

(Daniel Göring)