Bayesian Unfolding of H.E.S.S. energy spectra

Masterarbeit aus der Physik

Vorgelegt von

Maximilian Oberndörfer

13.04.2017

Erlangen Centre for Astroparticle Physics
Friedrich-Alexander-Universität
Erlangen-Nürnberg

Erster Gutachter: Prof. Dr. Christopher van Eldik
Zweiter Gutachter: Prof. Dr. Uli Katz

Abstract

The High Energetic Stereoscopic System (H.E.S.S.) measures cosmic particles by detecting Cherenkov light emitted from electromagnetic showers in the earth's atmosphere. During measurement H.E.S.S. has to deal with a non–vanishing energy resolution and an energy bias. In order to take these effects into account an iterative Bayesian Unfolding algorithm is implemented into the Heidelberg Analysis Program (hap). The constraint of obtaining data without *a priori* assumptions suggests the application of a Bayesian Unfolding algorithm. The algorithm is tested on simple distributions, data–like simulations and on real data, namely from PKS 2155–304 and the Crab nebula. In this sense the unfolding algorithm is compared to an already implemented Forward Folding algorithm. The unfolding algorithm is found to retrieve superior results in all tested cases. The algorithm works within hap.

Contents

1. Motivation

The universe as of today still remains vastly unexplored. In order to shed more light on this enormous area several approaches in astrophysics have been made, one of these being gamma–ray astronomy. In gamma–ray astronomy sources which accelerate cosmic particles up to very high energies are observed by measuring highly energetic photons. These highly energetic photons, or gamma–rays, are created in interactions of these cosmic particles with interstellar gas in close proximity to the source.
The advantage that gamma–rays hold over other messenger particles is their lack of interaction with interstellar magnetic fields, which cause charged cosmic rays to be deflected from their point of origin. An explanatory scheme may be seen in figure 1.1.

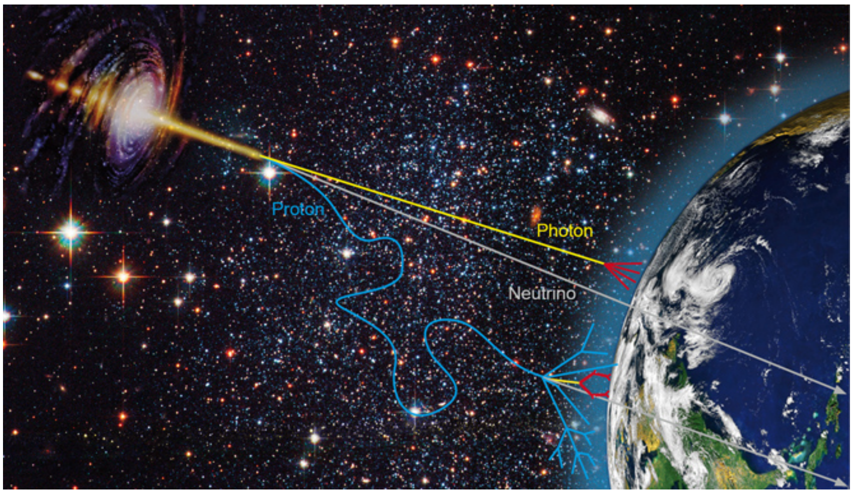

Figure 1.1: *Scheme of how messenger particles travel from their source to earth. While charged particles, such as protons, experience interactions with interstellar magentic fields, neutral particles, like photons and neutrinos, reach earth without being deflected. (Image taken from University of Nova Gorica, 2017).*

While the lack of deflection from the point of origin when observing gamma–rays is one point, the detection and measurement of gamma–rays is another, equally important one. This detection may either be taking place in space or on the surface of the earth.

Space–based telescopes are mounted on satellites and therefore limited in detection area and by weight. Their detection range is in the order of $\mathcal{O}(20\,\mathrm{MeV})$ to $\mathcal{O}(300\,\mathrm{GeV})$ (Funk, 2015). The most well known telescope of this kind is the Large Area Telescope (LAT) mounted on the Fermi satellite.

Ground–based telescopes on the other hand are not limited by their extent, but rather by suitable detection sites and financial means. A type of ground–based gamma–ray telescopes are the Imaging Atmospheric Cherenkov Telescopes (IACT). This kind of telescope detects Cherenkov light flashes from secondary particles produced by cosmic rays entering the atomosphere, e.g. photons. These secondary particles may reach a velocity higher than the speed of light in this medium, which causes them to emit Cherenkov light. For further explanation see figure 1.2.

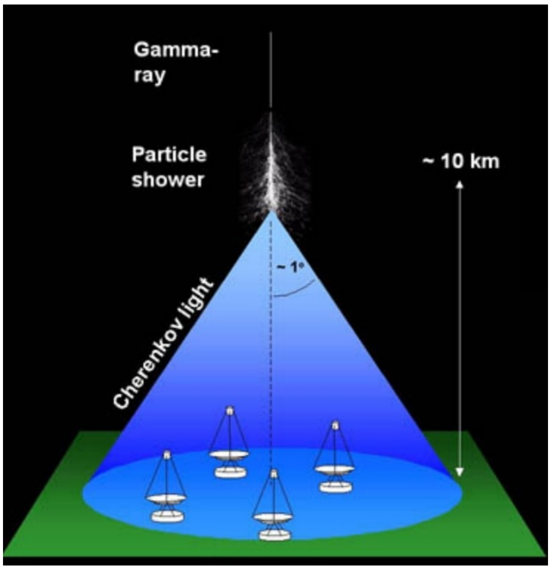

Figure 1.2: *The incident highly energetic photon undergoes pair production causing a shower of secondary particles, often referred to as cascade. This cascade due to its relativistic properties emits Cherenkov light at an angle of* $\approx 1\,^{\circ}$. *The Cherenkov light flash when reaching the surface of the earth lasts a short time interval in the order of* ns. *(Image taken from University of Nova Gorica, 2017).*

The requirement of having secondary particles moving at a speed greater than the speed of light and the Cherenkov flashes to be brighter than the background radiation limits the IACTs to the high energy range from $\mathcal{O}(30\,\mathrm{GeV})$ to $\mathcal{O}(100\,\mathrm{TeV})$ (Funk, 2015). The upper limit is set by statistics.

Due to the extended nature of Cherenkov flashes, typically in the order of $120\,\mathrm{m}^2$ at a 2000 m altitude (Funk, 2015), IACTs consist of several telescopes combined in an array, increasing the effective area required to measure the extended Cherenkov flashes. An example of an IACT is the High Energetic Stereoscopic System (H.E.S.S.), consisting out of five telescopes, located in Namibia.
Like any other instrument, H.E.S.S. also has underlying physical boundaries limiting the precision of the measurement. These effects may cause, for example, the reconstructed origin of a detected particle to be slightly off its true origin, or the reconstructed energy of a particle to be different from its true value.
In order to account for these effects, Monte–Carlo (MC) simulations are carried out, simulating and testing the average instrument behaviour under given conditions. The instrument behaviour changes over time due to the exposure of mirrors to sand or overall ageing of equipment causing a decrease in optical efficiency. Among other parameters

the instrument behaviour also depends on the zenith angle at which the observation took place. A great zenith angle corresponds to a longer path the shower has to propagate through the atmosphere to reach the telescope, complicating the analysis and reconstruction of the primary particle.
For the analysis of H.E.S.S. data these simulations are stored in lookups, having one lookup for one given set of parameters, e.g. the mentioned zenith angle and optical efficiency. Upon running the analysis chain the instrument response is derived from these lookups. An example of the instrument response when reconstructing energy, the energy migration matrix (EMM), may be seen in figure 1.3. Shown is an example of H.E.S.S. phase 2 data. H.E.S.S. has two optical phases. During the first phase only four telescopes were installed and used for measurements. After the addition of a fifth, larger telescope in the center of the already existing four telescopes in 2012 H.E.S.S. operates in phase 2.

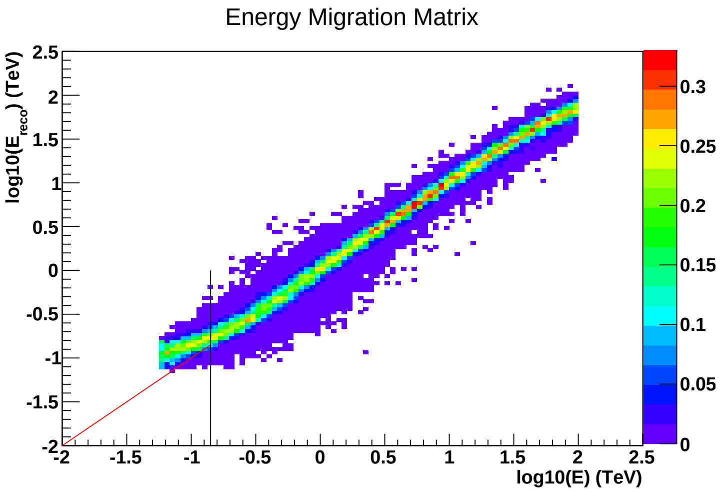

Figure 1.3: *An energy migration matrix (EMM) of H.E.S.S. phase 2 data of PKS 2155–304. The red line through origin indicates the path a non–biased matrix would follow. At about* 0.14 TeV, -0.85 *in the figure, indicated by the black line the energy bias exceeds a limit of* 10 % *off the line through origin.*

Under perfect conditions this matrix is diagonal, meaning every particle is reconstructed at exactly its true energy.
In reality not all particles of one given true energy are reconstructed at their precise true energy, but with a certain probability in an interval around it. The probability distribution here resembles a diagonal matrix smeared by a Gaussian distribution. This effect is called *energy resolution*.
Another effect disturbing the measurement is the so called *energy bias*. Particles affected by energy bias are not reconstructed at their true energy, but systematically at an energy shifted from the true energy value.
During measurement both effects occur, leading to a shift of the Gaussian distribution from the diagonal of the matrix.

In figure 1.3 a non–zero energy resolution may be seen throughout the entire range by the broadening of the matrix. Energy bias only plays a significant role at particle energies below 0.14 TeV, in the figure denoted by -0.85, seen by the shift from a roughly diagonal matrix towards a flat distribution on the left side of the plot.

In order to gain a deeper insight into energy spectra of astrophysical sources with H.E.S.S. one needs to relieve oneself of both effects.

One way to accomplish this is by using the method of *Forward Folding.* In Forward Folding an assumption about the source's spectrum is made and subsequently adjusted taking the EMM into account so that the estimate matches observation. The main restriction of this method is its necessity of an assumption about the spectral shape of the source.

Another way of relieving oneself of the effects of energy resolution and bias, which does not require any assumptions, is introduced in this thesis, the Bayesian Unfolding.

This thesis is organised as follows. The first chapter covers the theoretical background needed for applying unfolding in general and Bayesian Unfolding in particular. In the following chapter the Bayesian Unfolding algorithm is tested using simple and data–like simulations. Measurement and analysis methods with H.E.S.S. are explained, before data–driven simulations are introduced. In this way properties and boundaries of the algorithm are discussed. In the fourth chapter the application of the unfolding algorithm to real data, namely from the Crab nebula and the blazar PKS 2155–304, is presented. The conclusion summarises the so far obtained results and gives a short outlook on the future application of the unfolding algorithm.

2. Theoretical background of unfolding algorithms

In the first part of this chapter unfolding is defined and an application example is given. The second part introduces the Richardson–Lucy algorithm which motivated the Bayesian Unfolding algorithm implemented during this thesis. The latter algorithm is introduced in the third part of this chapter.

2.1 Definition *Unfolding*

The Oxford Dictionary (2017) defines unfolding, or *deconvolution*, as “The improvement of resolution of images or other data by a mathematical algorithm designed to separate the information from artefacts which result from the method of collecting it”. In this work precisely what is defined here is done.

In order to gain a deeper insight into unfolding, in the following the definition will be explained using the example of Fourier Deconvolution.
The following analogy depicts the principle behind Fourier Deconvolution.

> Fourier Deconvolution is the converse of Fourier convolution in the sense that division is the converse of multiplication. (O'Haver, 2008)

Indeed, Fourier Deconvolution is conducted using a point–by–point division of two signals in the Fourier domain and a subsequent reverse transformation to the original domain, e.g. time domain.
In practice one may use Fourier Deconvolution on following equation

$$h = f * g \tag{2.1}$$

in which f is a desired, unknown signal, g a known (or estimated) convolution signal, for example an instrument response function, and h the known measured signal.
f may be calculated by

$$F = H/G \tag{2.2}$$

$$f = F^{-1} \tag{2.3}$$

with F, G and H being the Fourier transforms of f, g and h.
An example of an actual application of Fourier Deconvolution can be seen in figure 2.1 (O'Haver, 2008).

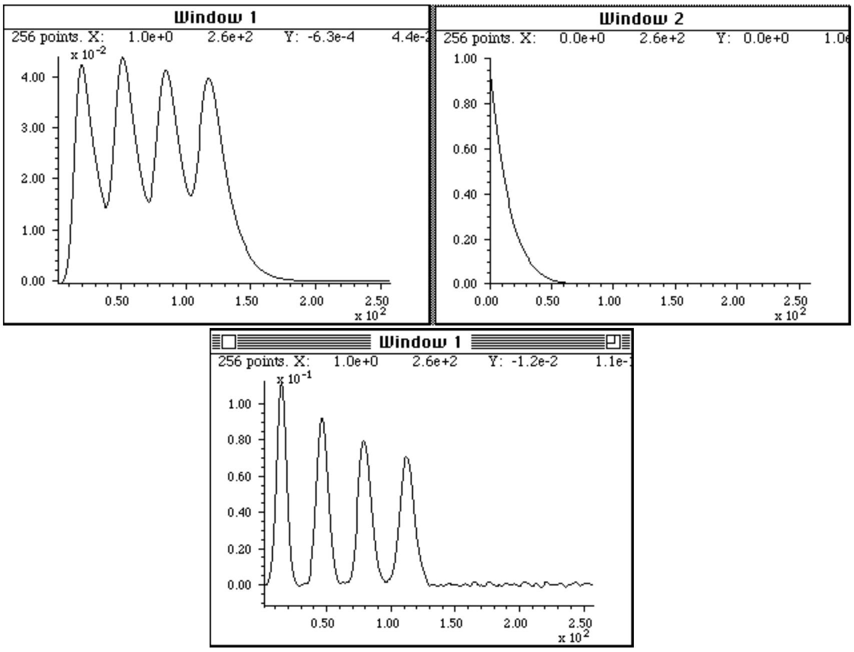

Figure 2.1: *The measured signal h (top left panel) has been recorded using a low-pass filter with an exponential tailing response function g (top right panel). The desired signal f (bottom panel) is then derived by deconvoluting the response function from the measured signal. (Image taken from O'Haver, 2008).*

2.2 Richardson-Lucy algorithm

In some cases applying Fourier Deconvolution will not lead to the desired result. One of these cases is described by W. H. Richardson in his paper introducing a *Bayesian–Based Iterative Method of Image Restoration* (Richardson, 1972). In his work Richardson focuses on the restoration of images, which have been distorted, for example, by a point spread function of a camera. He states that Fourier Deconvolution for these images only yields recognizable results as long as the noise level in the picture is very low, hence the need for a new unfolding algorithm.

In 1974 L. B. Lucy independently proposes *An iterative technique for the rectification of observed distributions* (Lucy, 1974), describing the same algorithm.

This algorithm is called the Richardson-Lucy algorithm.

In the following the Richardson-Lucy algorithm is introduced using the approach of W. H. Richardson, and its application explained using examples from both Richardson and Lucy.

Assuming H to be a degraded, recorded image, F to be the original image and G being the instrument response function, in this case a point spread function, one yields following equation

$$H = F * G \tag{2.4}$$

similar to equation (2.1). F,G, and H are assumed to be discrete probability–frequency functions, meaning the value of a point in one of these functions describes the frequency at which an event occurs at this point. For simplicity reasons G is normalized.

Being given the recorded image H and the instrument response G one may use Bayes's theorem to obtain the conditional probability of an event at F_i given an event at H_k as follows.

$$P(F_i|H_k) = \frac{P(H_k|F_i)P(F_i)}{\sum\limits_j P(H_k|F_j)P(F_j)} \tag{2.5}$$

with $i, j \in [1; I]$ and $k \in [1; K]$, I the size of the array F and K the size of the array H. Using $P(F_i|H_k) = P(F_iH_k)/P(H_k)$ and considering all H_k and their dependence on F_i according to G one yields

$$P(F_i) = \sum_k P(F_iH_k) = \sum_k P(F_i|H_k)P(H_k) \tag{2.6}$$

Replacing the conditional probability $P(F_i|H_k)$ from equation (2.5) in equation (2.6) one yields the following.

$$P(F_i) = \sum_k \frac{P(H_k|F_i)P(F_i)P(H_k)}{\sum\limits_j P(H_k|F_j)P(F_j)} \tag{2.7}$$

The problem at hand is $P(F_i)$ in the right side of the equation being the desired solution. Due to F_i being unknown one may estimate $P(F_i)$ in the right hand side of equation (2.7). In order to justify this estimate one may use an iterative process as described in equation (2.8).

$$P_{r+1}(F_i) = P_r(F_i) \sum_k \frac{P(H_k|F_i)P(H_k)}{\sum\limits_j P(H_k|F_j)P_r(F_j)} \tag{2.8}$$

with $r = 0, 1, ...$ and $P_0(F_i)$ being estimated.
$P_0(F_i)$ is assumed to be a uniform distribution of the form $P_0(F_i) = 1/I$ or $F_{i,0} = \sum\limits_i F_i/I$.
By replacing the probabilities in equation (2.8) in the following way

$$\begin{aligned} P(F_i) &= F_i/\sum_i F_i \\ P(H_k) &= H_k/\sum_k H_k \\ P(H_k|F_i) &= P(G_{i,k}) = G_{i,k} \end{aligned} \tag{2.9}$$

with $\sum\limits_k G_{i,k} = 1$, and making use of unfolding being a conservative process in the sense that $\sum\limits_i F_i = \sum\limits_k H_k$ the equation becomes the following.

$$\frac{F_{i,r+1}}{\sum_l F_l} = \frac{F_{i,r}}{\sum\limits_l F_l} \sum_k \frac{G_{i,k} \cdot (H_k/\sum\limits_l F_l)}{\sum\limits_j G_{j,k} \cdot (F_{j,r}/\sum\limits_l F_l)} \tag{2.10}$$

with $l \in [1; I]$, or

$$F_{i,r+1} = F_{i,r} \sum_k \frac{G_{i,k}H_k}{\sum\limits_j G_{j,k}F_{j,r}} \tag{2.11}$$

Equation (2.11) is the Richardson-Lucy algorithm. Convergence of the algorithm is shown by both Richardson (1972) and Lucy (1974).

In his work Lucy states that the algorithm essentially removes large scale deviations of F_r from the true F in one iteration. Further iterations only yield small corrections of F_r, which slowly match the statistical fluctuations in the measured sample H. This last property is used in section 2.3 in order to obtain an abort criterion.

In their respective works Richardson and Lucy present the working principle of their algorithm using examples. While Richardson focuses on the unfolding and restoration of images, Lucy applies his algorithm to astronomical data, e.g. measurements of stellar rotation in certain star classes. In the following, examples drawn from both works are introduced.

Starting with examples of image restoration, one may see the unfolding of degraded images with and without additional noise in figure 2.2.

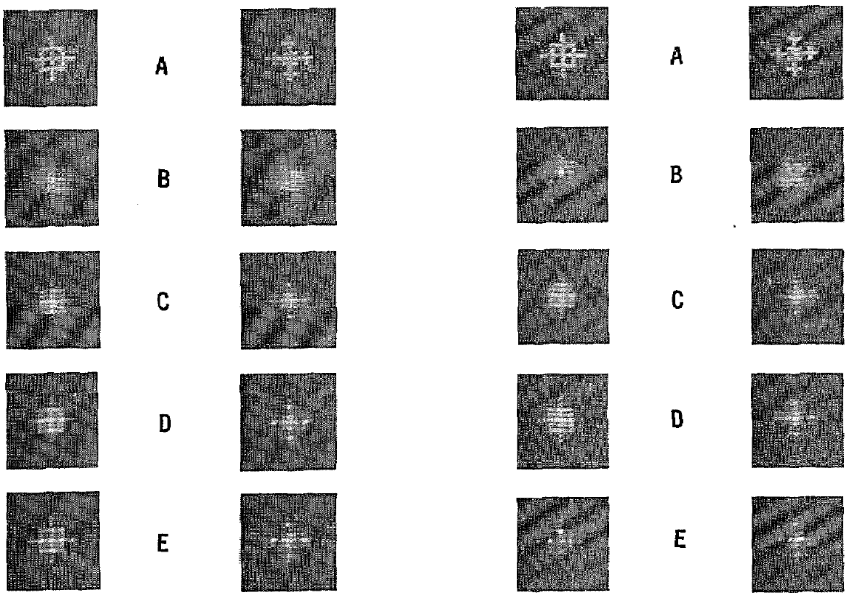

Figure 2.2: *Unfolding of two images without noise (left) and with an added noise (right). (A) shows the original image, (B) the degraded image, (C) the unfolded image after* 10 *iterations, (D) after* 20 *iterations and (E) after* 30 *iterations. The degradation was achieved by applying a uniform point spread function to the original image. The noise added to the right two images was of the form* $H' = H(1 + rd)$ *with* $r = 0.1$ *the noise level and* $d \in (0;1)$ *a random number. (Images taken from Richardson, 1972).*

Richardson states that after unfolding the degraded image for 10 iterations one may already identify the original image in both cases. In the case of no noise the unfolded image steadily improves with a rising number of iterations. In the case of added noise the unfolded image's quality seems to decline from 20 to 30 iterations, even though Richardson claims differently. This decline in quality might be explained by the unfolded data matching more and more the statistical fluctuations in the degraded image.
As initially stated, Fourier Deconvolution provides poor results when applied to images with background noise. In figure 2.3 a restoration of both example images with added noise using Fourier Deconvolution with an estimated least–square filter can be seen.
Even though the original images are reproduced when applying Fourier Deconvolution,

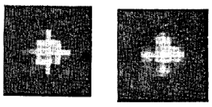

Figure 2.3: *Restoration of degraded images with additional $r = 0.1$ noise using Fourier Deconvolution with an estimated least–square filter. The presence of background noise is noticeable after unfolding. This background noise is not found in pictures unfolded using the Richardson–Lucy algorithm. (Image taken from Richardson, 1972).*

the unfolded images show a non–negligible amount of background noise, which may be a limiting factor when unfolding more complicated pictures. This background noise is not present in images unfolded using the Richardson–Lucy algorithm.

After having shown how applying the Richardson–Lucy algorithm to degraded images improves the image's quality heavily, one may also be interested in its application to measured, binned data in other fields of physics, for example astronomy.
The following example shows the unfolding of measured stellar rotation of 42 B6–B9e stars in figure 2.4.

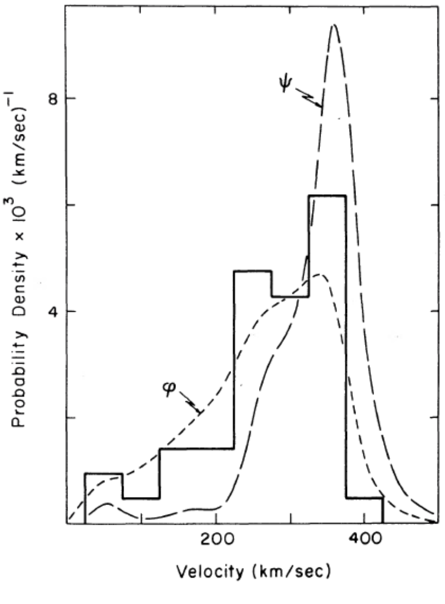

Figure 2.4: *Unfolding of $v \sin i$ measures for 42 B6–B9e stars. φ is a continuous function obtained from linear interpolation between the centers of each bin of raw data and has been taken as the input H into the unfolding algorithm. Ψ is the obtained result after applying unfolding for 4 iterations and shows a distinct peak at 360 km/s. (Image taken from Lucy, 1974).*

The figure shows a measurement of $v \sin i$ of these stars, which have been predicted to all have equatorial velocities of ≈ 350 km/s. After applying the Richardson-Lucy algorithm the distribution shows a sharp peek at 360 km/s, confirming the predictions.

2.3 Bayesian Unfolding of energy spectra

An algorithm similar to the Richardson–Lucy algorithm may be applied when unfolding energy spectra. The algorithm introduced here follows the one introduced by Wouters and Brun (D. Wouters, n.d.).
Before describing the algorithm used in this work, a change of notation is necessary.

The original and recorded data are no longer probability frequency functions F and H, but binned count histograms $N_{T,i}$ with $i \in [1; n_T]$ and $N_{R,j}$ with $j \in [1; n_R]$. These histograms are binned logarithmically equidistant in true energy E_T and reconstructed energy E_R with n_T and n_R being the total number of bins in true and reconstructed energy.
The instrument response function G, see equation (2.9), is now denoted by the energy migration matrix (EMM) $P(E_R|E_T)_{ij}$ with $i \in [1; n_T]$, $j \in [1; n_R]$. The EMM gives the probability of a particle with true energy $E_{T,i}$ to be detected in the reconstructed energy bin $E_{R,j}$ taking the finite energy resolution and energy bias of the instrument into account. It fulfills the following equation.

$$\sum_{j=1}^{n_R} P(E_R|E_T)_{ij} = 1 \quad \forall i$$

Unfolding of count spectra allows for a different physical interpretation of the unfolding process. For count spectra the following equation holds true.

$$N_{R,j} = \sum_{i=1}^{n_T} P(E_R|E_T)_{ij} N_{T,i} \tag{2.12}$$

Equation (2.12) implies that an inversion of the EMM and a subsequent multiplication by N_R yields the desired true energy spectrum N_T. Unfortunately direct inversion of the EMM leads to large errors on N_T due to small eigenvalues of $P(E_R|E_T)$, hence is not desirable.
Bayesian Unfolding offers a solution to this problem. Even though the EMM is not inverted, the true distribution N_T is approximated iteratively. The following equation shows the Richardson–Lucy algorithm (2.8) using the current notation

$$N_{T,i}^{(n+1)} = N_{T,i}^{(n)} \cdot \sum_{j}^{n_R} P(E_R|E_T)_{ij} \frac{N_{R,j}}{\sum_k^{n_T} P(E_R|E_T)_{kj} N_{T,k}^{(n)}} \tag{2.13}$$

with $n \in \mathbb{N}_0$ denoting the iteration. N_T^0 is estimated as a uniform distribution in the shape of

$$N_{T,i}^0 = \frac{\sum_{j=1}^{n_R} N_{R,j}}{n_R}$$

When measuring count spectra the number of counts in one bin is assumed to follow a Poisson distribution. In order to take these statistical fluctuations into account additional reconstructed mock spectra are simulated. A mock spectrum is generated by assigning random new numbers of counts for each bin within the Poisson error bars of the respective bin. The unfolding of all mock spectra separately and the subsequent averaging yields a stable solution (D. Wouters, n.d.). When unfolding with mock spectra, formula (2.13) becomes

$$N_{T,hi}^{(n+1)} = N_{T,hi}^{(n)} \cdot \sum_{j}^{n_R} P(E_R|E_T)_{ij} \frac{N_{R,hj}}{\sum_k^{n_T} P(E_R|E_T)_{kj} N_{T,hk}^{(n)}} \tag{2.14}$$

with $N_{T,hi}$ the $n_H \times n_T$ matrix and $N_{R,hj}$ the $n_H \times n_R$ matrix containing all mock spectra in true and reconstructed energy respectively, n_H being the number of simulated mock spectra.

Not only does this approach yield a more stable solution, but it also allows for the estimate of errors. In their note Wouters and Brun (D. Wouters, n.d.) propose the use of covariance matrices for this purpose.

In the following, $C_{R,jl}$ denotes the estimated covariance matrix of N_R with $j, l \in [1; n_R]$, and $C_{T,ik}$ the estimated covariance matrix of N_T with $i, k \in [1; n_T]$. The covariance matrix of the reconstructed count spectrum is estimated as shown in equation (2.15).

$$C_{R,ij} = \sum_{h=1}^{n_H} (N_{R,hi} - N_{R,i}) \cdot (N_{R,hj} - N_{R,j})/n_H \tag{2.15}$$

with $N_{R,i}$, $N_{R,j}$ denoting bin content of bins in the originally measured distribution, therefore being the known means of the reconstructed mock spectra.

The covariance matrix of the unfolded count spectrum is estimated in an analogous way, but the mean of the mock spectra distribution in true energy is estimated.

$$C^{(n)}_{T,ij} = \sum_{h=1}^{n_H} (N^{(n)}_{T,hi} - \overline{N}_{T,i}) \cdot (N^{(n)}_{T,hj} - \overline{N}_{T,j})/(n_H - 1) \tag{2.16}$$

with n denoting the iteration and $\overline{N}_{T,i} = \sum_{h=1}^{n_H} \frac{N^{(n)}_{T,hi}}{n_H}$.

The data measured in bins in reconstructed energy is uncorrelated, which yields a diagonal covariance matrix C_R. The elements in this diagonal matrix display the variance of each bin, therefore the square of the bin's standard deviation. Still assuming a Poisson distribution of counts in each bin the standard deviation for a large enough choice of n_H approaches

$$\sigma_{R,j} = \sqrt{N_{R,j}}$$

with $j \in [1; n_R]$. An example of an estimated covariance matrix C_R may be seen in figure 2.5.

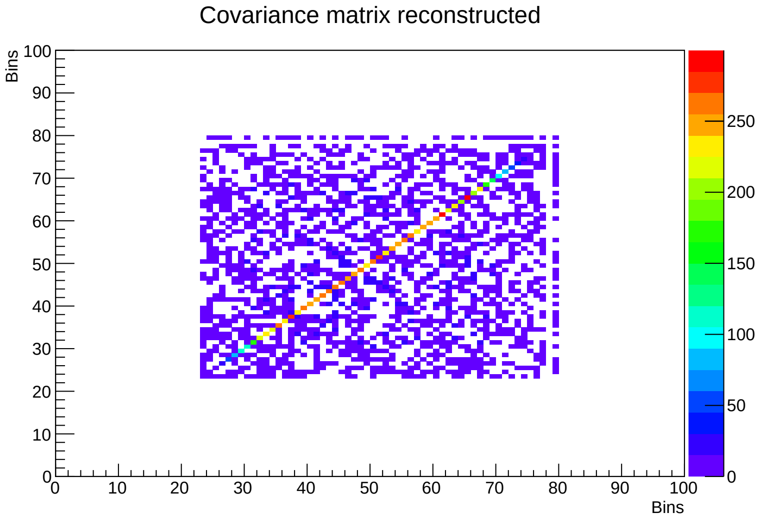

Figure 2.5: *An estimated covariance matrix* C_R *simulating* $n_H = 500$ *mock spectra of a true block distribution defined in the bin range* $[30; 70]$*. In an optimal case (*$n_H \to \infty$*) the off–diagonal elements are* 0*.*

During the unfolding process the data in the unfolded distribution becomes correlated, leading to a non–diagonal covariance matrix, as may be seen in figure 2.6.

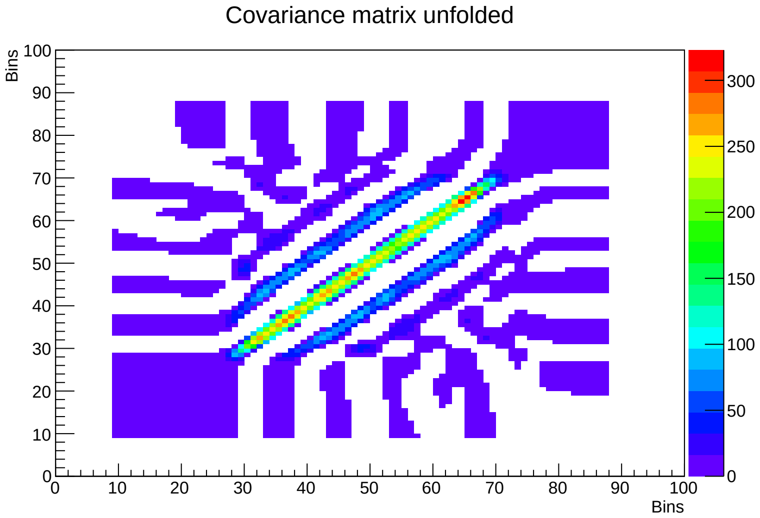

Figure 2.6: *An estimated covariance matrix* $C_T^{(20)}$ *simulating* $n_H = 500$ *mock spectra of a true block distribution defined in the bin range* $[30; 70]$ *smeared by a Gaussian distribution with* $\sigma = 0.3$ *and subsequent unfolding. The large non–zero off diagonal values indicate correlation between bins induced by the unfolding process.*

The variance of each bin in true energy, the diagonal elements of the covariance matrix C_T, correspond to the bin's statistical errors.

One may now use the covariance matrices of both reconstructed and unfolded distribution to obtain an abort criterion for the iterative unfolding algorithm. This abort criterion is received by demanding the traces of both matrices to be equal. Physically interpreted this demand means a conservation of statistical errors during unfolding (D. Wouters, n.d.). Since each mock spectrum is unfolded separately, the unfolded distributions deviate more with rising number of iteration, leading to a continuously increasing $\mathrm{tr}\left(C_T^{(n)}\right)$. The abort criterion then may be formulated as follows

$$\mathrm{tr}\left(C_T^{(n)}\right) / \,\mathrm{tr}(C_R) \geq 1 \qquad (2.17)$$

This abort criterion is tested thoroughly in chapter 3.

The examples shown in figure 2.5 and 2.6 are produced assuming $n_R = n_T$. The unfolding algorithm does not necessarily preserve the binning, but may rebin the unfolded distribution during unfolding. In order to reduce correlation one may choose a binning in the form of $n_T = n_R/1.4$. The value of 1.4 has been proposed as an optimal parameter by Albert et al. (2007). The rebinning during unfolding is also tested in chapter 3.

3. Simulation studies of the Bayesian Unfolding algorithm

In this chapter the Bayesian Unfolding algorithm introduced in section 2.3 is tested using simulations. The first section deals with simple unfolding problems. The algorithm is applied to simple mathematically defined functions and its behaviour for different sets of parameters is discussed. Furthermore the abort criterion is tested. In the second section unfolding of data–driven simulations is conducted and its results interpreted.

3.1 Unfolding of simple distributions

In this section the unfolding of Gaussian, block and halfcircle distributions smeared with a Gaussian distribution is conducted and properties of the unfolding algorithm are derived. The abort criterion in equation (2.17) is tested.

3.1.1 Generation and unfolding of simple distributions

In this subsection the simulation of simple distributions, which are used to test the unfolding algorithm, is explained. At first a mathematically defined normalised function $f_{\text{true}}(x)$ is created. In this section the following functions are used.

$$\begin{aligned} f_{\text{Block}}(x) &= 1/N \\ f_{\text{Gauss}}(x) &= \frac{1}{\sqrt{2\pi \cdot \sigma_{Gauss}^2}} \cdot e^{-\frac{x^2}{2 \cdot \sigma_{Gauss}^2}} \\ f_{\text{HC}}(x) &= \frac{1}{N} \cdot \sqrt{r^2 - x^2} \end{aligned}$$

with N being a suitable normalisation, for example $N = 4$ for $f_{\text{Block}}(x)$ with $x \in [-2; 2]$. According to the choice of function a histogram H_{true} is filled with n_{events} counts. This histogram subsequently is convolved with a Gaussian function of the form

$$f_{\text{IR}}(x) = \frac{1}{\sqrt{2\pi \cdot \sigma_{\text{res}}^2}} \cdot e^{-\frac{(x-b)^2}{2 \cdot \sigma_{\text{res}}^2}} \tag{3.1}$$

This function mimics the instrument response with a resolution σ_{res} and a bias b. The migration matrix, which is used for unfolding, is derived from equation (3.1). An example of a migration matrix may be seen in figure 3.1.
By convolving H_{true} with $f_{\text{IR}}(x)$ a histogram H_{IR} is obtained. In order to mimic the

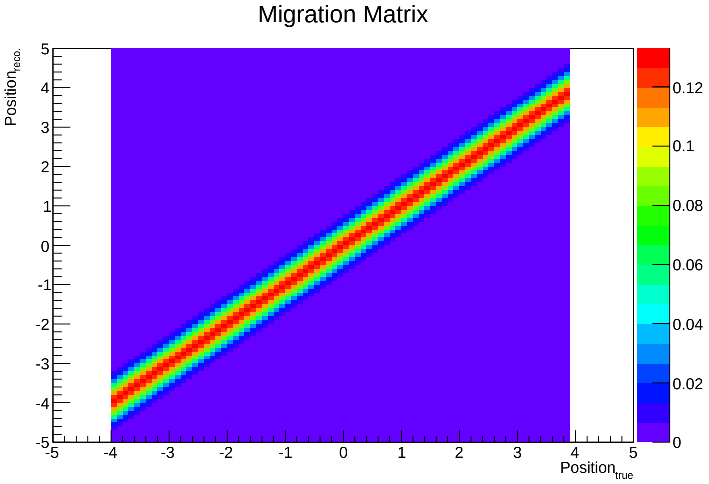

Figure 3.1: *A migration matrix derived from an instrument response function* $f_{IR}(x)$ *with resolution* $\sigma_{res} = 0.3$ *and bias* $b = 0$*. The binwidth is* 0.1*. The matrix is not filled in between* $[-5;-4]$ *and* $[3.9;5]$ *since the associated* H_{IR} *is* 0 *in these ranges and a reduced computing time is achieved in this way.*

observation the histogram H_{reco} is obtained dicing out n_{events} random events according to H_{IR}.

In order to reconstruct the original H_{true} the unfolding algorithm is applied to H_{reco}. Note that the unfolding is carried out for n_{H} mock histograms as described in section 2.3. The averaged result is denoted by H_{unf}.

The following examples show H_{true}, H_{reco} and H_{unf} for different choices of original distributions $f_{\mathrm{true}}(x)$. The simulations are carried out choosing $n_{\mathrm{H}} = 500$, $n_{\mathrm{events}} = 10000$, $\sigma_{\mathrm{res}} = 0.3$ and $b = 0$. The binwidth of all histograms has been set to 0.1. Figures 3.2, 3.3 and 3.4 show the unfolding results of the original distributions f_{Block}, f_{Gauss} and f_{HC}.

In figure 3.2 the unfolding of $f_{\mathrm{true}}(x) = f_{\mathrm{Block}}(x)$ with $x \in [-2;2]$ and $N = 4$ shows the algorithm's response to sharp edges. At the edges overshoots are created, followed by steep downfalls. This behaviour may lead to an insufficient unfolding result, since the overshoot differs heavily from the true distribution.

When choosing $f_{\mathrm{true}}(x) = f_{\mathrm{Gauss}}(x)$, seen in figure 3.3, the unfolded distribution matches the original distribution well. This matching may be due to the similarity of convolution function and true function.

The unfolding when choosing $f_{\mathrm{true}}(x) = f_{\mathrm{HC}}(x)$ with $r = \sqrt{2}$ and $N = \pi$, seen in figure 3.4, again introduces overshoots. The edges of the halfcircle seem too steep for the unfolding algorithm, yielding a behaviour analogous to $f_{\mathrm{true}}(x) = f_{\mathrm{Block}}(x)$. In figure 3.5 one may see the result of unfolding when choosing $f_{\mathrm{true}}(x) = f_{HC}(x)$ with radius $r = 2$ and $N = 2\pi$. The reduction of overshoots is clearly visible.

3.1.2 Testing the unfolding algorithm

In this subsection the unfolding algorithm is tested using a large number of simulated distributions. The maximum amount of iterations needed to fulfill the abort criterion in

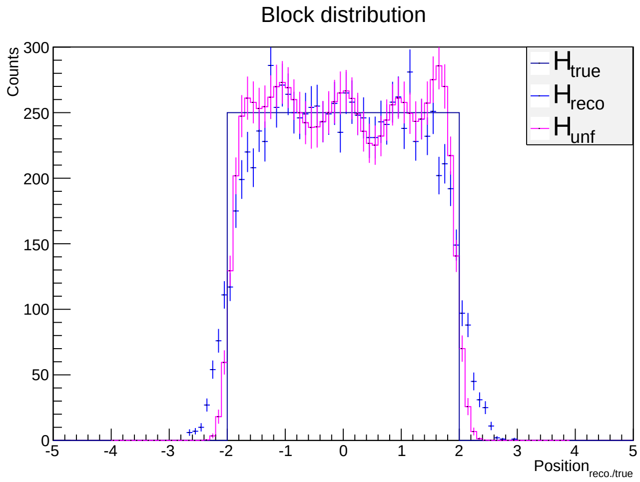

Figure 3.2: *$f_{true}(x)$ was chosen to be $f_{Block}(x)$ with $x \in [-2;2]$ and $N = 4$. The simulation is carried out with $n_H = 500$, $n_{events} = 10000$, $\sigma_{res} = 0.3$ and $b = 0$, iterating the unfolding algorithm 22 times. The error bars are the standard deviations obtained when deriving the covariance matrices.*

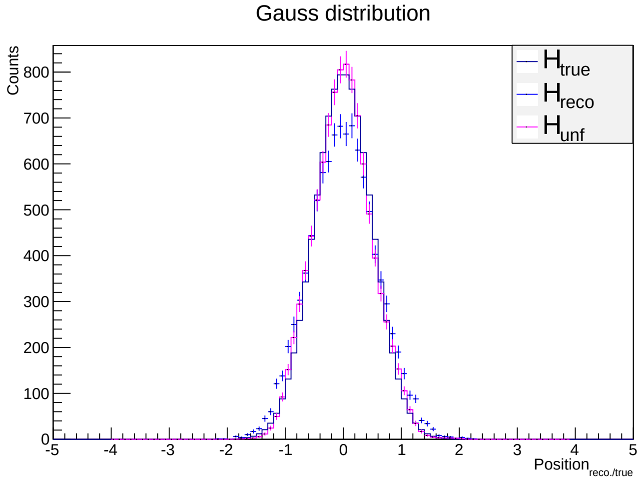

Figure 3.3: *$f_{true}(x)$ was chosen to be $f_{Gauss}(x)$ with $\sigma_{Gauss} = 0.5$. The simulation is carried out with $n_H = 500$, $n_{events} = 10000$, $\sigma_{res} = 0.3$ and $b = 0$, iterating the unfolding algorithm 22 times. The error bars are the standard deviations obtained when deriving the covariance matrices.*

dependency on the choice of $f_{\text{true}}(x)$, n_{events}, σ_{res} and b is determined. Additionally the quality of reconstructing the original distribution is tested using the Kolmogorov-Smirnov test. At last it is determined whether rebinning during unfolding yields an improvement in the sense of less correlation in the unfolded distribution and comparable quality in reconstructing the original distribution.
In order to provide stable results 500 simulations have been carried out for each set of parameters, simulating an additional $n_H = 500$ mock spectra within each simulation.

Before the results are discussed, the choices of parameters are shown exemplarily. The first example, seen in figure 3.6, shows different numbers of events n_{events}.

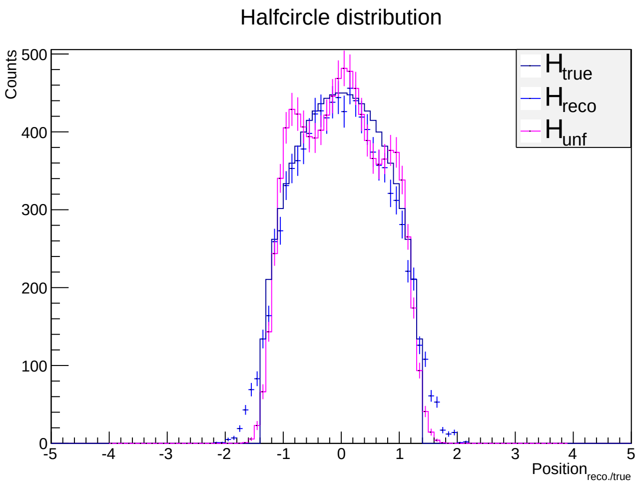

Figure 3.4: *$f_{true}(x)$ was chosen to be $f_{HC}(x)$ with $r = \sqrt{2}$ and $N = \pi$. The simulation is carried out choosing $n_H = 500$, $n_{events} = 10000$, $\sigma_{res} = 0.3$ and $b = 0$, iterating the unfolding algorithm 20 times. The error bars are the standard deviations obtained when deriving the covariance matrices.*

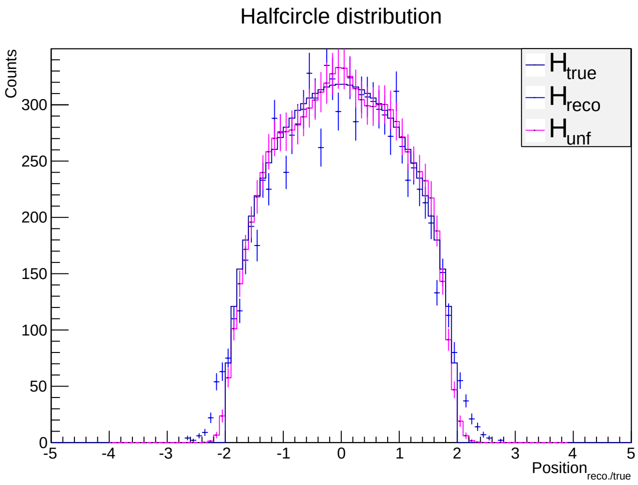

Figure 3.5: *Unfolded distribution with $f_{true}(x) = f_{HC}(x)$ choosing $r = 2$ and $N = 2\pi$. The simulation is carried out with $n_H = 500$, $n_{events} = 10000$, $\sigma_{res} = 0.3$ and $b = 0$, iterating the unfolding algorithm 22 times. The reduction of overshoots in comparison to figure 3.4 is clearly visible.*

Since the reconstructed mock distributions are diced out following a Poisson distribution for each bin, in the case of high statistics the reconstructed and unfolded distribution match the true distribution better in the sense of relative distance. The second example, seen in figure 3.7, displays the effect of different σ_{res}.

A greater distortion or worse resolution leads to a broadening of the reconstructed distribution. The third example, shown in figure 3.8, shows the effect a bias has on the reconstructed distribution.

A bias shifts the whole reconstructed distribution in regard to the true distribution. The unfolding algorithm is supposed to reverse this effect.

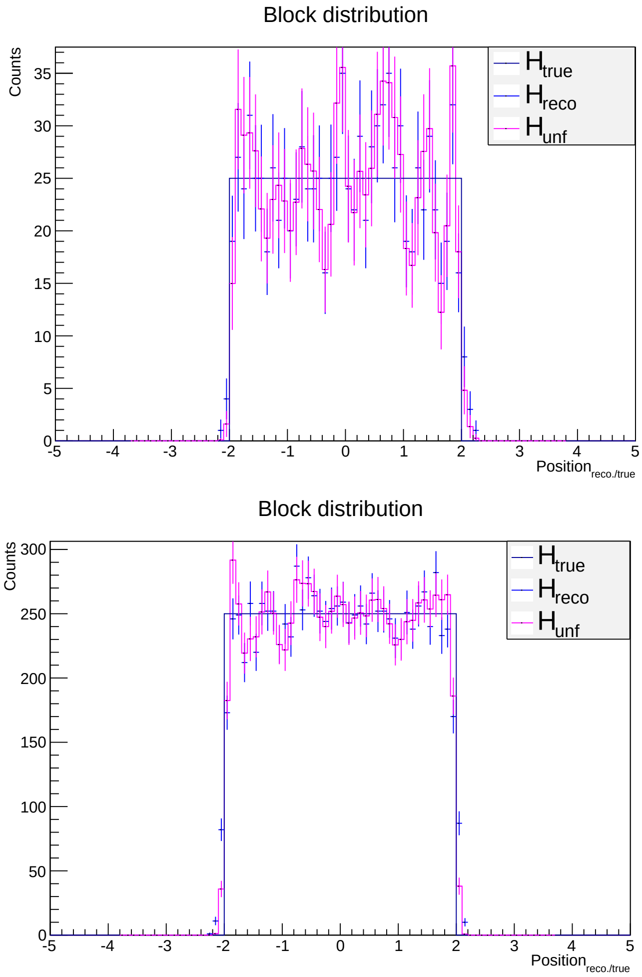

Figure 3.6: $n_{events} = 1000$ *(top) and* $n_{events} = 10000$ *(bottom). The other parameters are* $f_{true}(x) = f_{Block}(x)$ *with* $x \in [-2;2]$ *and* $N = 4\,,\ \sigma_{res} = 0.1$ *and* $b = 0\,.$

Testing iterations until abortion

The first test examines the dependency of the number of iterations needed to fulfill the abort criterion on different choices of $f_{\text{true}}(x)$, n_{events}, σ_{res} and b. In table 3.1 the results of this test may be seen.

The data in this table may be interpreted as follows. A small distortion of the measured data ($\sigma_{\text{res}} = 0.1$) leads to a small amount of iterations until the abort criterion is met. With increasing distortion also the number of iterations rises. An increase by a factor of 10 in distortion ($\sigma_{\text{res}} = 1.0$) yields an increase in number of iterations by a factor of 11.6 to 13.75. When choosing $f_{true} = f_{Gauss}$ fewer iterations are needed to fulfill the abort criterion in comparison to the other choices of f_{true}. The smaller amount of iterations needed may be explained by the similarity of f_{Gauss} to the convolution function, which also is a Gaussian function. The numbers of iterations needed to fulfill the abort critertion are equal within errors for the choices of $f_{true} = f_{Block}$ and $f_{true} = f_{HC}$, even though a slightly smaller number is needed in the latter, smoother case. The effect of energy bias b does not significantly affect the number of iterations needed to reach the abort criterion.

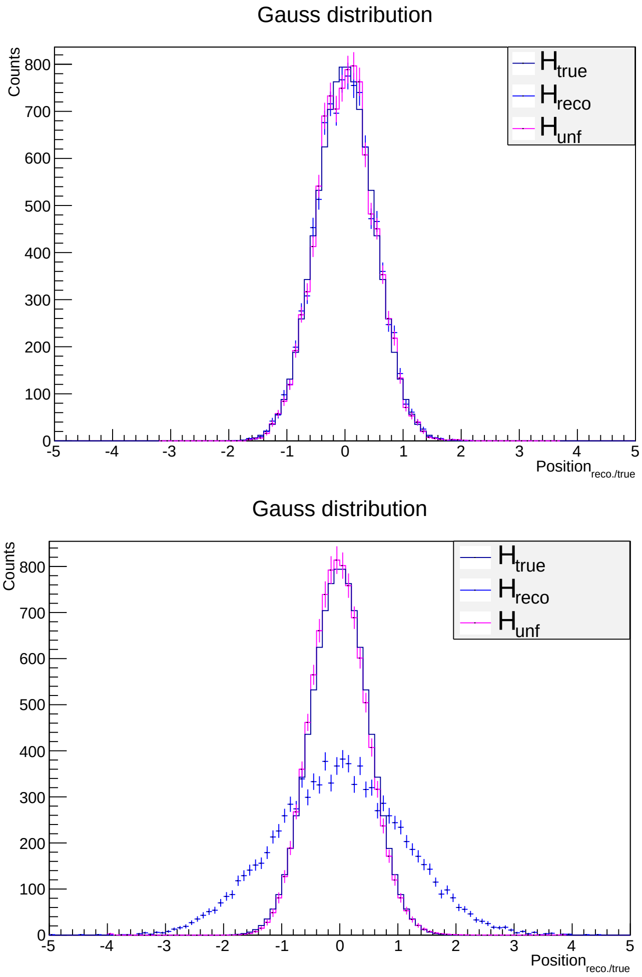

Figure 3.7: *$\sigma_{res} = 0.1$ (top) and $\sigma_{res} = 1.0$ (bottom). The other parameters are $f_{true}(x) = f_{Gauss}(x)$ with $\sigma_{Gauss} = 0.5$, $n_{events} = 10000$ and $b = 0$.*

σ_{res}	n_{events}	f_{Block}		f_{Gauss}		f_{HC}	
		b = 0	b = 1.0	b = 0	b = 1.0	b = 0	b = 1.0
0.1	1000	$6.00_{\pm 10^{-14}}$	$6.00_{\pm 10^{-14}}$	$5.99_{\pm 0.08}$	$6.00_{\pm 10^{-14}}$	$6.00_{\pm 10^{-14}}$	$6.00_{\pm 10^{-14}}$
	10000	$6.00_{\pm 10^{-14}}$	$6.00_{\pm 0.04}$	$5.98_{\pm 0.13}$	$5.99_{\pm 0.10}$	$5.98_{\pm 0.13}$	$5.99_{\pm 0.11}$
0.5	1000	$38.62_{\pm 1.19}$	$38.68_{\pm 1.31}$	$36.63_{\pm 1.46}$	$36.52_{\pm 1.46}$	$38.52_{\pm 1.24}$	$38.60_{\pm 1.30}$
	10000	$37.98_{\pm 1.20}$	$37.99_{\pm 1.20}$	$36.11_{\pm 1.35}$	$36.11_{\pm 1.37}$	$37.96_{\pm 1.31}$	$38.00_{\pm 1.23}$
1.0	1000	$80.54_{\pm 3.35}$	$82.48_{\pm 3.53}$	$70.85_{\pm 3.12}$	$71.11_{\pm 3.15}$	$78.98_{\pm 4.82}$	$80.77_{\pm 3.63}$
	10000	$79.30_{\pm 3.00}$	$81.27_{\pm 3.00}$	$69.54_{\pm 2.79}$	$69.65_{\pm 2.89}$	$77.76_{\pm 2.98}$	$79.50_{\pm 3.24}$

Table 3.1: *Number of iterations needed to fulfill the abort criterion. For each set of parameters 500 simulations have been carried out, each with $n_H = 500$ mock spectra. The displayed values are averaged over all simulations. The errors are the standard deviations.*

This last result aligns with Lucy's statement (Lucy, 1974) that the unfolding algorithm quickly deals with large scale distortions. Also the choice of statistics n_{events} does not

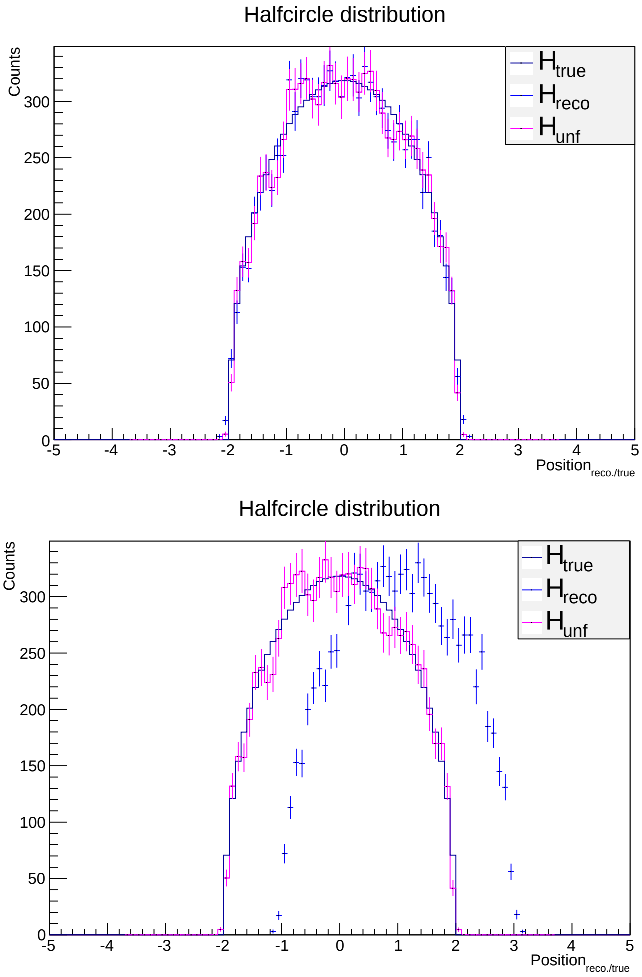

Figure 3.8: *$b = 0$ (top) and $b = 1.0$ (bottom). The other parameters are $f_{true}(x) = f_{HC}(x)$ with $r = 2$ and $N = 2\pi$, $\sigma_{res} = 0.1$ and $n_{events} = 10000$.*

influence the amount of iterations significantly.
Even though a smaller number of iterations needed to fulfill the abort criterion is desirable in the sense of computing time, it by definition means a larger variance of mock spectra per iteration. The question is whether the results provided by less iterations are worse in quality or not. This question is answered in the following paragraphs.

The Kolmogorv-Smirnov test

As initally stated, the quality of reconstructing the original distribution is tested by using the Kolmogorv-Smirnov (K–S) test. In the next passage the K–S test is introduced following the approach of J. Hedderich (2016).
The test is used to determine whether measured data was drawn from a specific distribution. It relies on the empirical distribution function, which is given for measurements $x_{(1)}$, $x_{(2)}$, ..., $x_{(n)}$, given in increasing order, by

$$F_x(x_{(i)}) = \frac{n(i)}{n} = \frac{i}{n}$$

with $n(i)$ the amount of measurements smaller than x_i.
The K–S test is based on the maximum distance of the empirical distribution function to a chosen cumulative distribution function F_0. The empirical distribution function F_x is tested on this chosen distribution function. The K–S test is not dependent on the choice of F_0, yet following conditions must be fulfilled. It may only be applied to continuous distributions. F_0 has to be fully defined, e.g. in expectancy value and variance, otherwise the test result is not accurate. The K–S test checks the following hypotheses

H_0: the data follows a distribution function F_0
H_A: the data does not follow a distribution F_0

The Kolmogorov-Smirnov statistic is given by

$$D = \sup_x |F_0(x_{(i)}) - F_x(x_{(i)})|. \tag{3.2}$$

In figure 3.9 one may see an example of a model distribution function F_0 (continuous line) and a measured distribution function F_x (histogram).

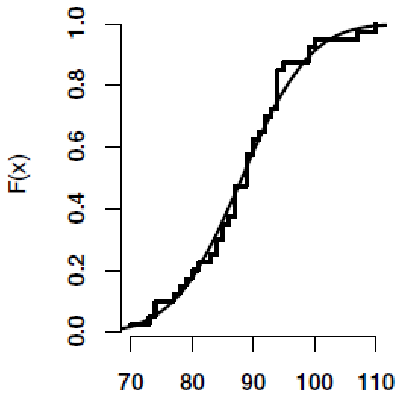

Figure 3.9: *Example of a theoretical distribution function F_0 (continuous line) and a measured distribution function F_x (histogram). D is determined by calculating the distance between each bin of the measured distribution and the theoretical distribution function and then taking its maximum. In this example the measured distribution of fasting blood sugar F_x is tested on a normal distribution F_0. (Image taken from J. Hedderich (2016))*

Testing the quality of unfolding results

The actual K-S test is conducted by comparing the distance D yielded from comparing observation to model with the critical values of the Kolmogorov distribution. The values of the Kolmogorov distribution may be obtained from tables or by using the following approximation formula:

$$D_\alpha = \sqrt{\frac{-0.5 \cdot \ln(\alpha/2)}{n}} \tag{3.3}$$

with n being the measured statistics and α the significance level which the null hypothesis is tested on. If D is larger than D_α the null hypothesis is rejected at a level α.
In the following tests a significance level of $\alpha = 0.05$ is used. When conducting many tests, this means that the null hypothesis should be rejected for a maximum of 5 % of

tests, if the measured distributions were drawn from the theoretical distribution. If the measured distributions were not drawn from the theoretical distribution, the amount of rejected null hypotheses should be higher than 5 %.

In this sense not only the quality of reconstructing the original distribution is tested, but it is also determined, whether aborting the algorithm even for $\mathrm{tr}\left(C_T^{(n)}\right) / \mathrm{tr}(C_R) < 1$ yields better results. In particular the cases

$$|\,\mathrm{tr}\left(C_T^{(n-1)}\right) / \mathrm{tr}(C_R) - 1| < |\,\mathrm{tr}\left(C_T^{(n)}\right) / \mathrm{tr}(C_R) - 1| \tag{3.4}$$

are of interest, with n the number of iterations needed to fulfill the abort criterion.

The distributions at test are the unfolded distributions of D in the final iteration for different choices of parameters and their respective reconstructed counterparts. D is derived as described above, testing the reconstructed and unfolded distribution on the true distribution. An example of such distributions of D may be seen in figure 3.10

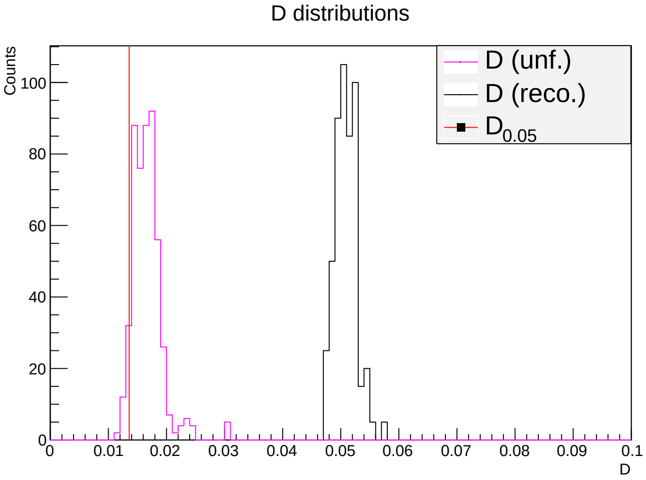

Figure 3.10: *Example of the distributions of D for the reconstructed (black) and unfolded distributions (pink). The red line indicates $D_{0.05}$. If either distribution has 95 % of its entries left of the red line, the distribution is said to be in agreement with the test distribution at a significance level of 5 %, hence the closer the distribution is to 0 the better the agreement. The distribution tested on is the originally simulated true distribution. The parameters chosen for this simulation were $f_{true} = f_{Block}$, $\sigma_{res} = 0.5$, $n_{events} = 10000$ and $b = 0$.*

The results of this test without taking bias into account are displayed in table 3.2.
The following conclusions may be drawn from this table. When applying unfolding to distributions which only experienced a small distortion, no significant improvement is achieved. Instead the data becomes correlated during unfolding. When dealing with small distortions and no bias unfolding hence does not yield desirable results. With increasing distortion the unfolded distribution always is in better agreement with the true distribution than the reconstructed one. The choice of n_{events} also plays an important role. With greater statistics the limit $D_{0.05}$ decreases according to equation (3.3). The improvement achieved with unfolding does not seem to match this decrease. The earlier mentioned smaller number of iterations needed to reach the abort criterion does not

σ_{res}	n_{events}	f_{Block}		f_{Gauss}		f_{HC}	
		unf.	reco.	unf.	reco.	unf.	reco.
0.1	1000	92.8 %	96.0 %	94.0 %	95.0 %	94.2 %	97.0 %
	10000	97.0 %	96.0 %	97.4 %	95.0 %	97.0 %	97.0 %
0.5	1000	78.4 %	2.0 %	73.2 %	0 %	80.6 %	78.0 %
	10000	4.4 %	0 %	78.6 %	0 %	70.2 %	0 %
1.0	1000	35.6 %	0 %	54.8 %	0 %	47.9 %	0.2 %
	10000	0 %	0 %	47.0 %	0 %	12.6 %	0 %

Table 3.2: *Percentage of D contained within the 5 % significance level when choosing $b = 0$. The reconstructed (reco.) or unfolded (unf.) distribution is in agreement with the true distribution if at maximum 5 % are not contained within the significance limit $D_{0.05}$. For each set of parameters 500 simulations have been carried out, each with $n_H = 500$ mock spectra.*

correspond to worse results. The distributions which have been subject to great distortion and required a great number of iterations until the abort criterion was met match the true distribution less than the distributions which have been unfolded with less iterations. The quality of reconstruction heavily depends on the distortion of the measured distribution. An important conclusion is that apart from the case of small distortion, in which unfolding is not desired, the unfolding algorithm fails to produce distributions which are found to be in agreement with the original one according to the K–S test. Even though unfolding does not reproduce the original distribution, it still yields an improvement in comparison to the reconstructed distribution for large distortions. In the following this improvement is examined in the presence of bias.

When simulating bias $b = 1.0$, one yields the results shown in table 3.3.

σ_{res}	n_{events}	f_{Block}		f_{Gauss}		f_{HC}	
		unf.	reco.	unf.	reco.	unf.	reco.
0.1	1000	93.0 %	0 %	94.0 %	0 %	93.2 %	0 %
	10000	97.4 %	0 %	99.8 %	0 %	96.0 %	0 %
0.5	1000	78.8 %	0 %	72.8 %	0 %	81.6 %	0 %
	10000	5.2 %	0 %	75.0 %	0 %	72.0 %	0 %
1.0	1000	36.8 %	0 %	54.0 %	0 %	46.4 %	0 %
	10000	0 %	0 %	61.2 %	0 %	6.6 %	0 %

Table 3.3: *Percentage of D contained within the 5 % significance level when choosing $b = 1.0$. The reconstructed (reco.) or unfolded (unf.) distribution is in agreement with the true distribution if at maximum 5 % are not contained within the significance level. For each set of parameters 500 simulations have been carried out, each with $n_H = 500$ mock spectra.*

In the presence of bias the reconstructed distributions even for small distortions are not in agreement with the original distribution. In the case of bias unfolding yields superior results and is hence desired regardless of the resolution. The results obtained for the unfolded distributions when simulating bias and the ones without bias are in agreement.

In both tables 3.2 and 3.3 the values when choosing $f_{true} = f_{Block}$, $\sigma_{res} = 1.0$ and $n_{events} = 10000$ are $0\,\%$. In this case also an improvement is achieved by applying unfolding, yet the distributions of D do not have any entries below $D_{0.05}$. Both distributions may be seen in figure 3.11.

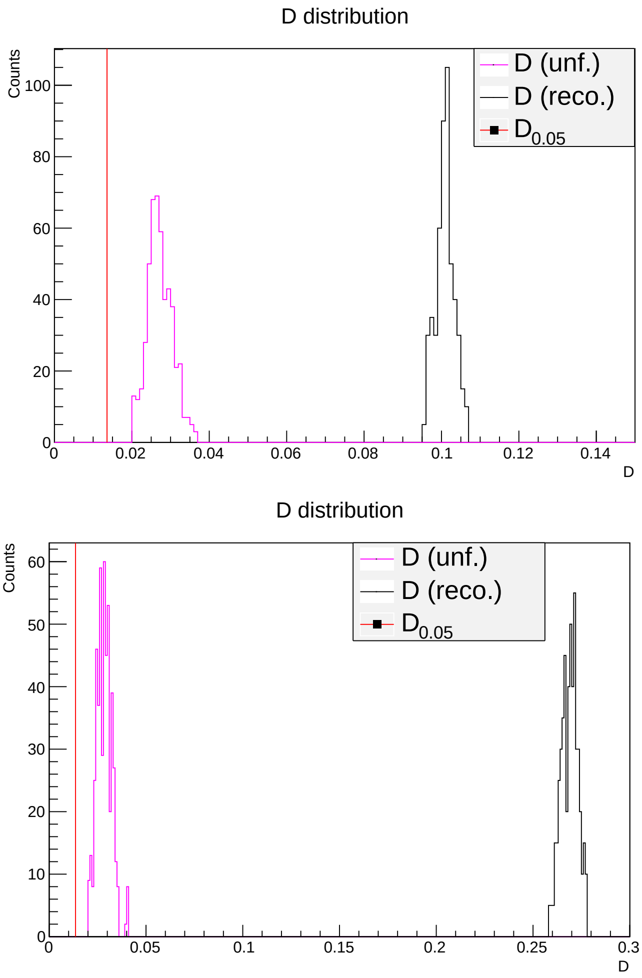

Figure 3.11: *Distributions of D for choosing $f_{true} = f_{Block}$, $\sigma_{res} = 1.0$, $n_{events} = 10000$ and bias $b = 0$ (top) and $b = 1.0$ (bottom). An improvement in both cases may be seen by the shift of the unfolded distribution (pink) towards 0 in regard to the reconstructed distribution (black).*

The next test determines whether an abortion of the unfolding algorithm even if

$$\mathrm{tr}\left(C_T^{(n)}\right) / \,\mathrm{tr}(C_R) < 1$$

yields better results if equation (3.4) holds true. In the following table 3.4 the difference of percentages of D

$$\Delta_{(n),(n-1)} = \frac{\int_0^{D_{0.05}} D^{(n)}}{5} - \frac{\int_0^{D_{0.05}} D^{(n-1)}}{5} \tag{3.5}$$

contained within the $5\,\%$ significance limit of the final iteration (n) and the iteration prior to the final one (n-1) without bias is displayed. The factor of 5 in the denominator is due to the display of 500 simulations in units of $\%$.

σ_{res}	n_{events}	f_{Block}	f_{Gauss}	f_{HC}
0.1	1000	$-0.4\,\%$	$0\,\%$	$-0.6\,\%$
	10000	$0\,\%$	$0\,\%$	$0\,\%$
0.5	1000	$0\,\%$	$-0.4\,\%$	$-0.6\,\%$
	10000	$-0.2\,\%$	$-0.4\,\%$	$0\,\%$
1.0	1000	$0.2\,\%$	$0\,\%$	$-0.6\,\%$
	10000	$0\,\%$	$-0.4\,\%$	$0\,\%$

Table 3.4: $\Delta_{(n),(n-1)}$ *of the percentages of* D *contained within* $D_{0.05}$ *when choosing* $b = 0$ *of the* (n)*-th and* $(n-1)$*-th iteration. The differences are negligible.*

The differences obtained are negligible. The abort criterion remains as described in equation (2.17). The same holds true in the presence of bias.

Testing rebinning during unfolding

The last test will determine whether rebinning during unfolding is desirable or not. In a first approach the amount of iterations needed to fulfill the abort critertion when rebinning during unfolding is applied is subject to test. The rebinning applied here follows the one proposed by Albert et al. (2007) in the form of $n_T = n_R/1.4$. The numbers of iterations needed when applying rebinning and without rebinning, both without simulating bias, are shown in table 3.5.

σ_{res}	n_{events}	f_{Block}		f_{Gauss}		f_{HC}	
		rebin	normal	rebin	normal	rebin	normal
0.1	1000	$4.00_{\pm 10^{-15}}$	$6.00_{\pm 10^{-14}}$	$4.00_{\pm 10^{-15}}$	$5.99_{\pm 0.08}$	$4.00_{\pm 10^{-15}}$	$6.00_{\pm 10^{-14}}$
	10000	$4.00_{\pm 10^{-15}}$	$6.00_{\pm 10^{-14}}$	$4.00_{\pm 10^{-15}}$	$5.98_{\pm 0.13}$	$4.00_{\pm 10^{-15}}$	$5.98_{\pm 0.13}$
0.5	1000	$26.31_{\pm 0.84}$	$38.62_{\pm 1.19}$	$25.23_{\pm 0.98}$	$36.63_{\pm 1.46}$	$26.28_{\pm 0.87}$	$38.52_{\pm 1.24}$
	10000	$26.03_{\pm 0.86}$	$37.98_{\pm 1.20}$	$25.12_{\pm 1.02}$	$36.11_{\pm 1.35}$	$26.02_{\pm 0.94}$	$37.96_{\pm 1.31}$
1.0	1000	$55.80_{\pm 2.26}$	$80.54_{\pm 3.35}$	$50.38_{\pm 2.08}$	$70.85_{\pm 3.12}$	$54.79_{\pm 2.37}$	$79.98_{\pm 4.82}$
	10000	$55.42_{\pm 2.27}$	$79.30_{\pm 3.00}$	$49.82_{\pm 1.97}$	$69.54_{\pm 2.79}$	$54.29_{\pm 2.19}$	$77.76_{\pm 2.98}$

Table 3.5: *Number of iterations needed to fulfill the abort criterion comparing rebinning during unfolding (rebin) and no rebinning during unfolding (normal). For each set of parameters* 500 *simulations have been carried out, each with* $n_H = 500$ *mock spectra. The displayed values are averaged over all simulations. The errors are the standard deviations.*

The number of iterations needed to fulfill the abort criterion decreases if rebinning during unfolding is applied. This result is in accordance to expectation, since the wider binning causes greater values in the diagonal of the covariance matrix. The number of iterations is approximately decreased by the same factor the number of bins is decreased by. In the presence of bias comparable results are obtained.

Not only does the number of necessary iterations decrease, but also the correlation between bins in the unfolded distribution is reduced. The data corresponding to this statement can be seen in table 3.6. Displayed are the sums of absolute off–diagonal values of the covariance matrices $\sum\limits_{i\neq j}|C_{T,ij}|$ for applied rebinning (rebin) and no rebinning (normal) without simulating bias.

σ_{res}	n_{events}	f_{Block}		f_{Gauss}		f_{HC}	
		rebin	normal	rebin	normal	rebin	normal
0.1	1000	$2006_{\pm 64}$	$3465_{\pm 93}$	$1506_{\pm 65}$	$2820_{\pm 101}$	$1974_{\pm 62}$	$3435_{\pm 98}$
	10000	$20603_{\pm 610}$	$36161_{\pm 925}$	$15531_{\pm 625}$	$29390_{\pm 1109}$	$20255_{\pm 609}$	$35717_{\pm 1072}$
0.5	1000	$9582_{\pm 324}$	$13919_{\pm 489}$	$6722_{\pm 264}$	$9872_{\pm 373}$	$9385_{\pm 311}$	$13669_{\pm 404}$
	10000	$97064_{\pm 2693}$	$142103_{\pm 3604}$	$68260_{\pm 1854}$	$101644_{\pm 2338}$	$95316_{\pm 2562}$	$140015_{\pm 3443}$
1.0	1000	$15240_{\pm 617}$	$21345_{\pm 983}$	$9179_{\pm 627}$	$12886_{\pm 928}$	$14329_{\pm 539}$	$20180_{\pm 815}$
	10000	$153850_{\pm 3962}$	$218425_{\pm 5733}$	$92951_{\pm 2551}$	$131690_{\pm 3896}$	$145548_{\pm 3137}$	$207272_{\pm 4917}$

Table 3.6: $\sum\limits_{i\neq j}|C_{T,ij}|$ *for different choices of parameters comparing rebinning during unfolding (rebin) and no rebinning during unfolding (normal). For each set of parameters* 500 *simulations have been carried out, each with* $n_H = 500$ *mock spectra. The displayed values are averaged over all simulations. The errors are the standard deviations.*

As expected the covariances and hence the correlation between bins in the unfolded distribution decreases when applying rebinning. This result also holds true in the presence of bias. With a smaller number of iterations needed to reach the abort criterion and a decreased correlation between bins, rebinning seems to be the preferrable choice. In order to check whether it is really superior to unfolding without rebinning, the K–S test is also conducted for unfolded rebinned distributions. In table 3.7 the differences of D contained within $D_{0.05}$ defined as

$$\Delta_{rebin} = \frac{\int_0^{D_{0.05}} D}{5} - \frac{\int_0^{D_{0.05}} D_{rebin}}{5} \tag{3.6}$$

are shown, without simulating bias. The factor of 5 in the denominator is due to the display of 500 simulations in units of %.

σ_{res}	n_{events}	f_{Block}	f_{Gauss}	f_{HC}
0.1	1000	−1.8 %	0 %	−2 %
	10000	74.4 %	−0.6 %	0.2 %
0.5	1000	−5.8 %	−7.2 %	−4.6 %
	10000	4.4 %	−8.8 %	−0.8 %
1.0	1000	3 %	−6.6 %	−7.3 %
	10000	0 %	6.6 %	12.4 %

Table 3.7: Δ_{rebin} *of the percentages of* D *contained within* $D_{0.05}$ *when applying no rebinning and when applying rebinning during unfolding.*

The conclusions drawn from these results are that rebinning especially when simulating low statistics ($n_{events} = 1000$) provides comparable or slightly better restults. When simulating great statistics ($n_{events} = 10000$) and a small distortion ($\sigma_{res} = 0.1$), choosing a not smooth true function (f_{Block}), the option of rebinning proves to be the worse choice in comparison to no rebinning. Also when simulating great statistics ($n_{events} = 10000$), choosing smooth true functions (f_{Gauss}, f_{HC}) and a great distortion ($\sigma_{res} = 1.0$) unfolding with rebinning seems to deliver worse results.
The overall results do not provide enough evidence to justify using the unfolding algorithm solely with rebinning. Due to this lack of evidence data–driven simulations are tested for both rebinning and no rebinning during unfolding in the following section 3.2.

The main results of this subsection may be summarised as follows. Applying unfolding to measured data is only desirable if the data is severly distorted or biased in regard to the true data. The greater the distortion of the measured distribution, the more iterations are needed to fulfill the abort criterion. The unfolded distribution does not reconstruct the exact true distribution in the desired cases, but still yields an improvement in comparison to the reconstructed distribution. Rebinning during unfolding could prove to be superior to an equal binning, since less iterations are needed to fulfill the abort criterion and the data is less correlated when choosing wider binwidths, but a test of quality when rebinning does not allow for a final conclusion. Rebinning hence stays subject to test in the following sections.

3.2 Unfolding of data–driven simulations

In this section the unfolding of data–driven simulations is discussed. In the first subsection measurement and analysis methods with H.E.S.S. are introduced. This introduction gives a basic idea of how the unfolding algorithm works within the Heidelberg Analysis Program (*hap*) and how data–driven simulations should be carried out. The second subsection describes in detail how data–driven simulations are generated. In the third subsection the Bayesian Unfolding algorithm is applied to these simulations. In this way properties and boundaries of the algorithm when applied to data–like simulations are discussed.

3.2.1 Measurement and analysis methods with H.E.S.S.

In order to understand the generation of data–driven simulations and the problems occuring when unfolding real data, the measurement and analysis methods of H.E.S.S. have to be understood first. In this subsection these measurement and analysis methods are explained.

When measuring Cherenkov air showers with the H.E.S.S. telescopes, one yields images like the ones shown in figure 3.12.
Unfortunately these air showers are not only produced by the desired γ–rays but also by other cosmic rays, such as muons or protons. In order to separate the γ–like events from these background events it is sufficient to apply cuts on the shape of the air shower in the camera. While γ–like events show an elongated ellipse, seen in figure 3.12 (right), the shape of background events usually differs from this shape heavily. μ events for example generate a ring–like shape in the camera (see figure 3.12 (left)).

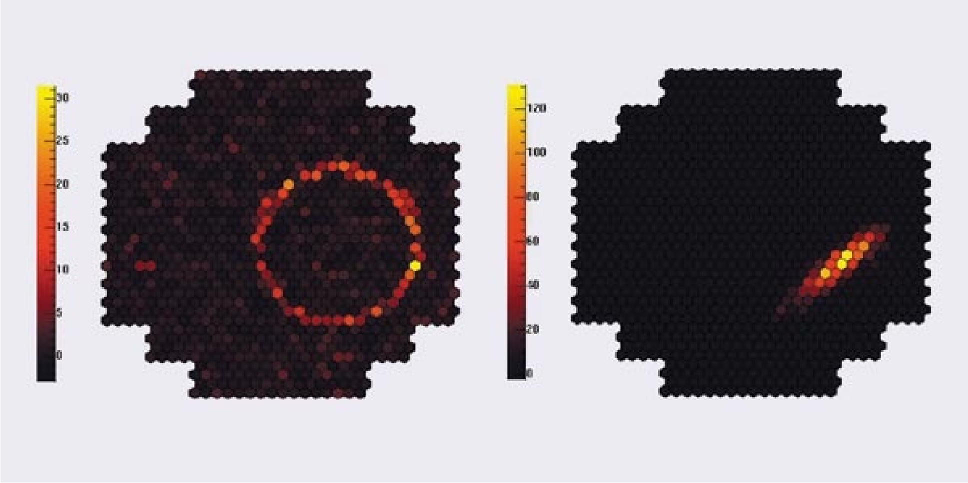

Figure 3.12: *Images of an Cherenkov air shower of a μ (left) and a γ–like particle (right) recorded with H.E.S.S. (Image taken from Hofmann, 2005).*

γ–like background events are caused by cosmic rays, e.g. protons, which during interaction in the atmosphere may create highly relativistic γ as secondary particles. These secondary particles induce Cherenkov air showers similar to the ones induced by primary photons. The separation of γ–like background events and γ–ray events may not be achieved by applying shape cuts. In order to subtract this background additionally to the region of the source, the ON region, several OFF regions, regions without any apparent source, are measured. The measurement of OFF regions is conducted simultaneously to the measurement of the ON region by choosing suitable regions within the field of view of the telescope. The background within the ON region then is estimated by multiplying the amount of OFF region counts N^{OFF} by a background normalisation factor α, which is defined as the ratio of the size of the ON region over the total size of OFF regions.

After having applied the shape cuts, one may now reconstruct the energy of the incoming γ–like particle.
In H.E.S.S. reconstruction of energies is achieved by using lookups. Lookups store parameters describing the performance of the instrument under given observation conditions (Gast, 2012). These parameters are generated prior to observations using Monte–Carlo simulations. For each set of observation parameters one lookup is created.
When reconstructing the initial energy of a γ–like particle from its detected shower, the parameters are optical efficiency, azimuth angle, zenith angle, offset from the observation position and telescope pattern.[1] In order to save computing time and limit data storage, the amount of sets is kept at a minimum and the range of parameters is discretised, e.g. in steps of 10° for zenith angles. Upon querying a certain set of parameters when running the analysis the lookups corresponding to the observation parameters closest to the queried set are retrieved and a linear interpolation is conducted (Gast, 2012).

In the case of energy reconstruction the lookup returns the energy that the initial γ–like particle most likely had under given observation conditions in dependence of the number

---

[1] H.E.S.S. may detect events with 2, 3 or 4 telescopes in phase 1 and 2, 3, 4 or 5 telescopes in phase 2. The telescope pattern is the combination of telescopes which detected the measured event.

of photo electrons detected in the camera of the telescope.

In this way energy count spectra of observed sources are generated. An example of such an energy count spectrum may be seen in figure 3.13.

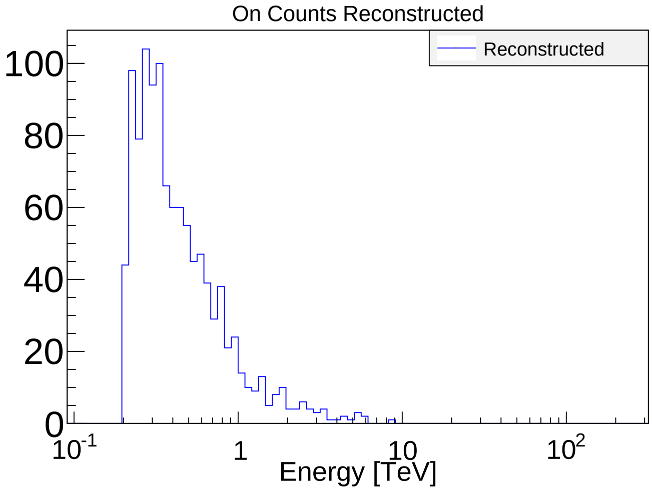

Figure 3.13: *Energy count spectrum of PKS 2155–304 from* 28 *runs with H.E.S.S. in phase 1. The energy is given in reconstructed energy E_{reco}. The y–axis is given in counts.*

The count spectra obtained in this way are affected by the instrument response of the telescopes. The x–axis of each spectrum therefore corresponds to a reconstructed energy E_{reco} instead of the original energy E_{true}. The reconstructed energy count spectrum is smeared by the energy resolution of the instrument and possibly even biased in regard to the true energy count spectrum, which complicates an analysis of the source's true count and flux spectrum.

The expected smearing and bias of a reconstructed spectrum is also stored in lookups. From these lookups the energy migration matrix (EMM) is derived during analysis. An example of an EMM may be seen in figure 3.14
The values stored in the EMM are the probabilities of a particle with true energy $E_{T,j}$ to be reconstructed in a bin of reconstructed energy $E_{R,i}$.

Count spectra are not suitable to reach conclusions about the source's spectrum yet. A more sophisticated result is obtained when examining flux spectra. For this purpose additional selections and cuts as well as the loss of events due to trigger conditions have to be taken into account. In *hap* these effects are stored in the effective area $A(E)$ lookups, which are available for reconstructed and true energy. The effective areas are given in units of cm^2.

A physical flux in units of $\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$ may be obtained from equation (3.7).

$$\left(\frac{dN}{dE}\right)_i = \frac{1}{\Delta E_i T_{obs} A(E_i)}(N_i^{\text{ON}} - \alpha N_i^{\text{OFF}}) \tag{3.7}$$

with ΔE_i the binwidth, T_{obs} the observation time and $A(E_i)$ the effective area in the chosen energy bin. Effective areas are given in true and reconstructed energy for ON and

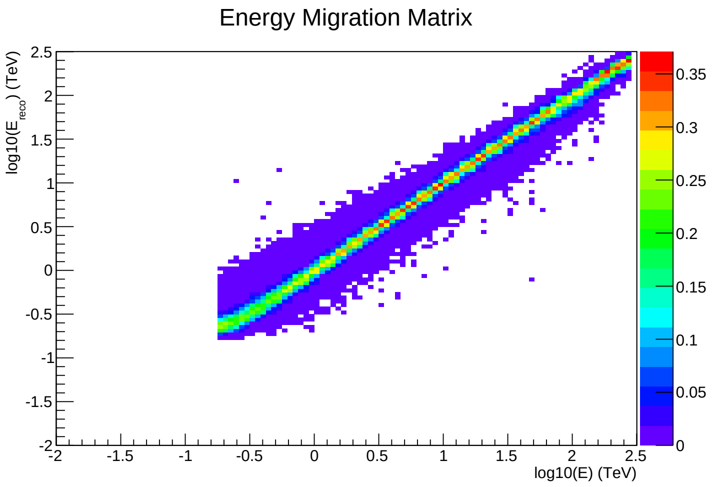

Figure 3.14: *Probabilites of particles with true energy $E_{T,j}$, j denoting the columns on the x–axis given in $\log_{10}(E_T)$, to be reconstructed in a reconstructed energy bin $E_{R,i}$, i denoting the rows on the y–axis given in $\log_{10}(E_R)$. Shown is an example of an energy migration matrix (EMM) of* 28 *runs of PKS-2155 304 with H.E.S.S. in phase 1.*

OFF regions separately.

When combining several measurements or runs, with one run corresponding to measuring 28 min, of one single source the quantities introduced above have to be combined correctly. In *hap* the combined effective areas are given by formula (3.8), turning the effective area A into effective area–time AT.

$$AT(E_i) = \sum_{r=1}^{n_{runs}} A(E_i)^r T^r \tag{3.8}$$

with T^r the respective observation time of run r and n_{runs} the amount of runs which are combined. This formula is caluclated for reconstructed and true energy.
The combined EMM is given by equation (3.9).

$$P(E_R|E_T)_{ij} = \frac{\sum_{r=1}^{n_{runs}} P(E_R|E_T)^r_{ij} A(E_{T,j})^r T^r}{AT(E_{T,j})} \tag{3.9}$$

The additional factor $A(E_{T,j})^r T^r$ ensures a correct weighing by effective area and observation time for each run.

In order to obtain the true energy spectrum from a reconstructed one, several methods may be applied making use of the EMM. The method described in this work is the unfolding of the reconstructed spectrum.
Another method is the *Forward Folding*. In Forward Folding a certain parametrization or shape of the true spectrum is assumed. The counts in each bin of reconstructed energy are estimated from start parameters taking into account the energy resolution and bias obtained from the EMM. The parameters then are adjusted so that the estimate matches the observation.

This method has proven to be very effective when the spectral shape of the source is well known. If such knowledge is lacking or if the analysis of a source's spectrum without assumptions is required, Forward Folding is not applicable. In this case the method of unfolding can be used.
Forward Folding has been implemented into *hap* prior to this work. A comparison of the unfolding and Forward Folding methods is provided in subsection 3.2.3.

3.2.2 Generation of data–driven simulations

In this subsection the generation of data–driven simulations is explained. As the name implies, data–driven simulations rely on quantities obtained from real data analysis. In detail, the EMM and effective areas in true and reconstructed energy are used here. In this simulation no background is simulated, hence all spectra are ON region spectra without an underlying γ–like background.

The first step in simulating data–driven simulations is the generation of a true flux spectrum in the shape of

$$\phi(E) = \phi_0 \cdot \left(\frac{E}{E_0}\right)^{-\Gamma} \tag{3.10}$$

with the normalisation ϕ_0 given in units of $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the spectral index Γ and $E_0 = 1\,\mathrm{TeV}$ without loss of generality throughout the entire thesis.
In order to obtain an energy count spectrum N_T^{ON} from this flux, the effective area–time $AT(E_T)$ has to be taken into account. The flux is integrated over the binwidth of the true effective area–time bin i and subsequently multiplied by $AT(E_{T,i})$.
From this count spectrum in true energy the count spectrum in reconstructed energy may be derived in the following way. The EMM stores the estimated distribution of counts from one true energy bin measured in several reconstructed energy bins. According to this distribution an amount of $N_{T,i}^{ON}$ events is diced out and filled in the reconstructed energy count spectrum N_R^{ON}. The total reconstructed energy count spectrum is obtained after having applied this method to all bins in true energy. The comparison of a true energy count spectrum and a reconstructed energy count spectrum may be seen in figure 3.15.
The reconstructed energy count spectrum N_R^{ON} is then unfolded using the Bayesian Unfolding algorithm with the same EMM that has been used to fill it, yielding an unfolded spectrum N_{unf}^{ON}. A comparison of a true, a reconstructed and an unfolded count spectrum may be seen in figure 3.16.
Since the original input is a flux spectrum, one wishes to reconstruct flux spectra again after unfolding in order to have a suitable comparison. Both, the reconstructed energy count spectrum and the unfolded energy count spectrum, are used to obtain flux spectra. After unfolding the unfolded count spectrum is given in true energy, hence $AT(E_T)$ has to be taken into account for the unfolded spectrum.
From measurement the exposure is only given in reconstructed energy. The exposure in true energy has to be estimated in this step. This estimate is obtained by normalizing the EMM for bins in reconstructed energy, yielding EMM_{exp}. The exposure in reconstructed energy then is distributed according to EMM_{exp} in true energy.

Before calculating flux spectra it may be mentioned that by filling bins in reconstructed energy only considering the EMM, bins in reconstructed energy may be filled, which have not been filled during real measurement and have an effective area–time $AT(E_R) = 0$.

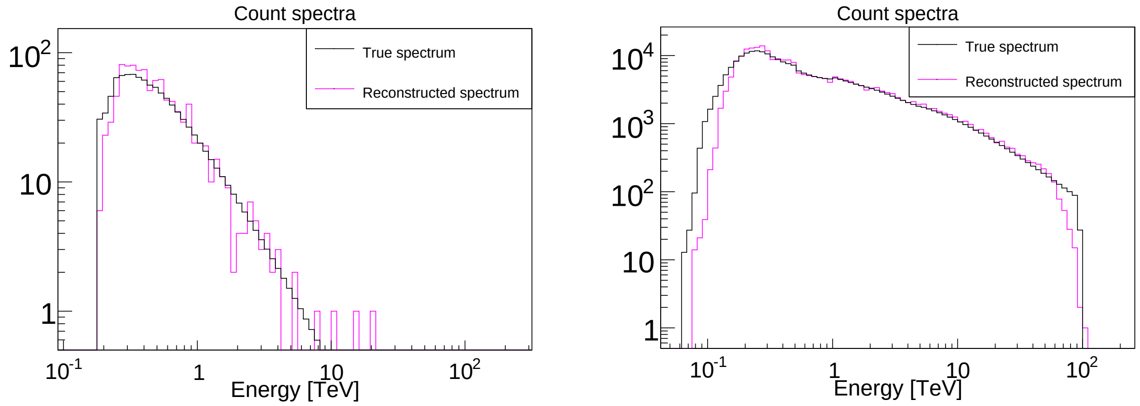

Figure 3.15: *Comparison of a true energy count spectrum and a reconstructed energy count spectrum. The underlying true energy flux is simulated with* $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, $\Gamma = 3$ *and data taken from* 28 *runs of PKS 2155–304 with H.E.S.S. in phase 1 (left). The true energy flux spectrum is simulated with* $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, $\Gamma = 2$ *and data taken from* 61 *runs of PKS 2155–304 with H.E.S.S. in phase 2 (right).*

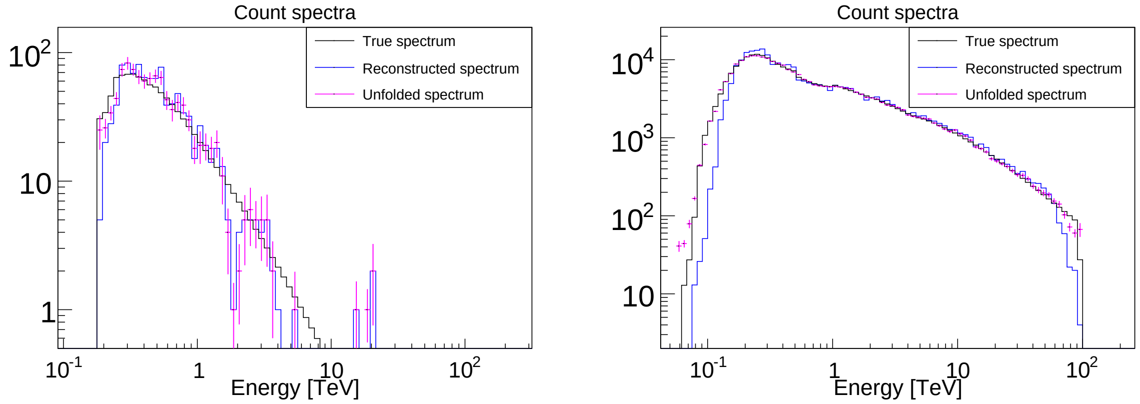

Figure 3.16: *Comparison of a true energy count spectrum, a reconstructed energy count spectrum and the unfolded count spectrum obtained after* 10 *iterations. The underlying true energy flux is simulated with* $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, $\Gamma = 3$ *and data taken from* 28 *runs of PKS 2155–304 with H.E.S.S. in phase 1 (left). The true energy flux spectrum is simulated with* $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, $\Gamma = 2$ *and data taken from* 61 *runs of PKS 2155–304 with H.E.S.S. in phase 2 (right).*

These bins are not taken into account when calculating flux spectra.

One yields the reconstructed and unfolded flux spectra by dividing the count spectra N^{ON} by their respective effective area–times, $AT(E_R)$ in the case of N_R^{ON} and $AT(E_T)$ in the case of N_T^{ON}.
These fluxes then are fitted with a function

$$\phi_{fit}(E) = \phi_{0,fit} \cdot \left(\frac{E}{E_0}\right)^{-\Gamma_{fit}} \tag{3.11}$$

and the obtained fit parameters are compared to the originally chosen parameters in equation (3.10). An example of such flux spectra and their respective fits may be seen in figure 3.17.

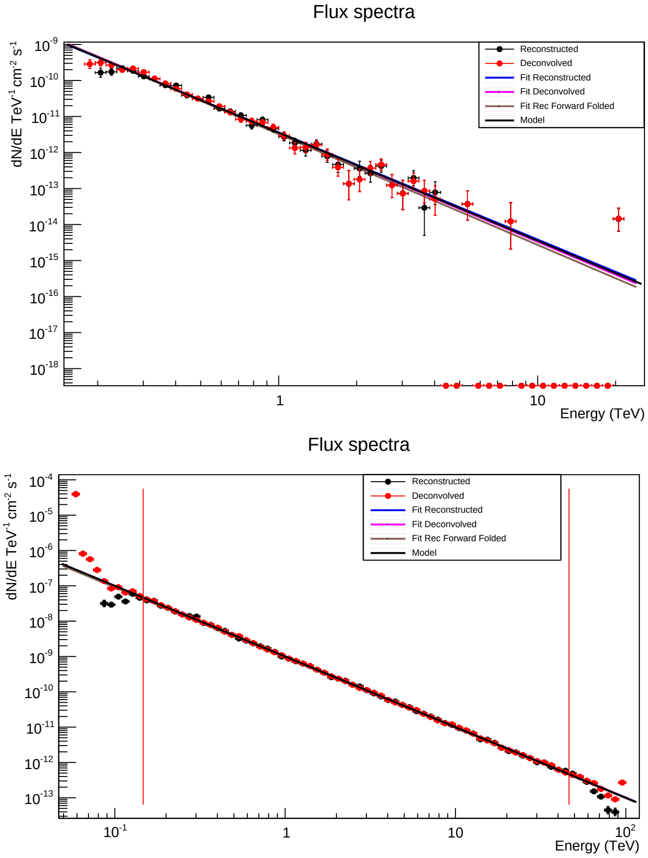

Figure 3.17: *Flux spectra of a simulated reconstructed count spectrum and the corresponding unfolded count spectrum, simulating a true flux spectrum with* $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ *and* $\Gamma = 3$ *using data from* 28 *runs of PKS 2155–304 with H.E.S.S. in phase 1 (top) and* $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ *and* $\Gamma = 2$ *using data from* 61 *runs of PKS 2155–304 with H.E.S.S. in phase 2 (bottom). The blue line indicates the flux fitted to* N_R^{ON} *in reconstructed energy using a log–likelihood fit, the brown line indicates the flux fitted to* N_R^{ON} *in true energy using a forward folded fit and the pink line indicates the flux fitted to* N_T^{ON} *in true energy using a log–likelihood fit. The black line shows the original flux spectrum in true energy. The vertical red lines in the bottom figure depict the safe energy range, the range in which all runs contribute to the combined spectrum.*

The unfolded flux spectrum seems to be unstable at the edges, showing great disparity between flux points and true flux spectrum. This effect is caused by small numbers of remaining counts in the count spectrum at these edges and small values of AT in these regions. Whether the effect disturbs the unfolded flux spectra is tested in the following subsection.

3.2.3 Unfolding of data–driven simulations

In this subsection the results of unfolding data–driven simulations is presented and interpreted. The quality of reconstructing count spectra is tested using the Kolmogorov–Smirnov (K–S) test as described in subsection 3.1.2. The reconstruction of original flux spectra is tested by fitting the resulting flux spectra in reconstructed and true energy and comparing the fit parameters to the input parameters. The simulations are carried out for both H.E.S.S. phase 1 and phase 2. In phase 1 the bias introduced by measurement is comparably small and the unfolded data is not supposed to be much superior to the reconstructed one. During measurements in phase 2 a non–negligible bias is introduced, in which case unfolding is supposed to deliver superior results like shown in subsection 3.1.2. The tests in phase 1 are carried out with and without rebinning during unfolding. Phase 2 is only tested without rebinning, since the necessary code has not been implemented into hap at this point of time.

The tests are carried out simulating true flux spectra with normalisations ϕ_0 in units of $10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and indices Γ with

$$\begin{aligned} \phi_0 &\in [0.356; 0.529; 100.0] \\ \Gamma &\in [2; 3] \end{aligned}$$

The arbitrary value of $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is testing the unfolding of flux spectra with very high statistics with a differential flux of 289.9 % of the Crab flux (Aharonian et al., 2008). Both indices are simulated choosing $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The smaller values of $0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $0.529 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ test the unfolding for realistic statistics of $\mathcal{O}(1000)$ events with 10.3 % and 15.3 % of the Crab flux (Aharonian et al., 2008). The index of 2 is simulated with a normalisation $\phi_0 = 0.529 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and the index of 3 is simulated choosing $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

In order to provide stable results each set of parameters is simulated about 10000 times, each time unfolding $n_H = 500$ mock spectra. As stated in the previous subsection the effect of unstable border regions in flux spectra is tested by fitting the flux spectra once including this range and once excluding it.

Simulations based on PKS 2155–304 measured with H.E.S.S. phase 1 without rebinning

At first the simulations taking the EMM and AT from 28 measurements of PKS 2155–304 with H.E.S.S. in phase 1 without applying rebinning during unfolding are discussed. The EMM and AT used may be seen in figure 3.18.

The EMM is rather diagonal and only introduces a minor bias at low energies. Table 3.8 shows the results of the K–S test of reconstructed and unfolded count spectra tested over the full energy range.

When simulating great statistics and choosing an index of $\Gamma = 3$ the distribution of D does not have any entries below $D_{0.05}$. In this case still an improvement is achieved by applying unfolding, as my be seen in figure 3.19.

The data in table 3.8 may be interpreted as follows. When simulating great statistics applying unfolding yields superior results for both choices of index. When simulating realistic statistics with an index of $\Gamma = 2$ the unfolded count spectrum seems to be slightly inferior to the reconstructed count spectrum. This may be caused by the harder spectrum choosing $\Gamma = 2$, since more particles at high energies are simulated. At high energies almost no energy bias is present, yielding a reconstructed spectrum which is in

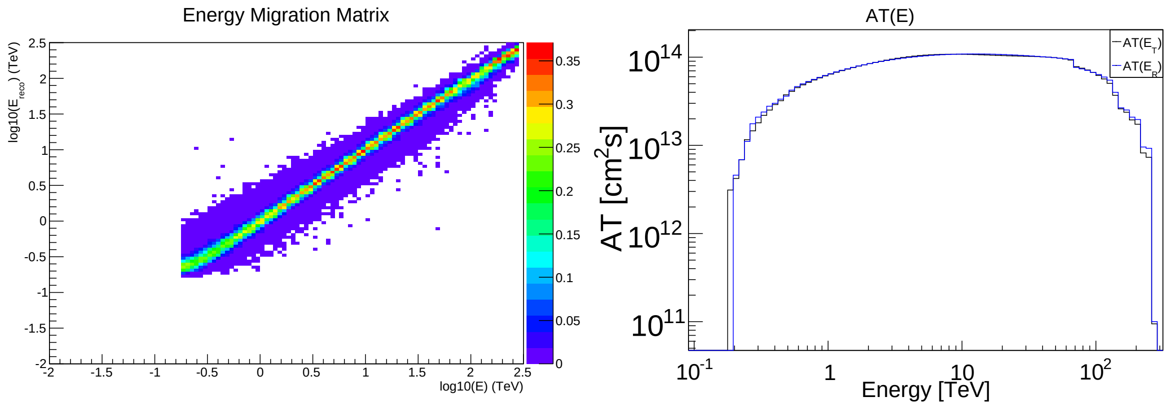

Figure 3.18: *Energy migration matrix (EMM) (left) and area–time (AT) (right) derived from analysing* 28 *measurements of PKS 2155–304 with H.E.S.S. in phase 1.*

ϕ_0 [TeV^{-1}cm^{-2}s^{-1}]	Γ	reco.	unf.
$100.0 \cdot 10^{-11}$	2	0 %	4.6 %
	3	0 %	0 %
$0.529 \cdot 10^{-11}$	2	78.6 %	69.6 %
$0.356 \cdot 10^{-11}$	3	14.3 %	64.0 %

Table 3.8: *Percentage of D contained within the* 5 % *significance level. The reconstructed (reco.) or unfolded (unf.) count spectrum is in agreement with the true distribution if at maximum* 5 % *are not contained within the significance limit* $D_{0.05}$*. The simulated count spectra are based on H.E.S.S. phase 1 data.*

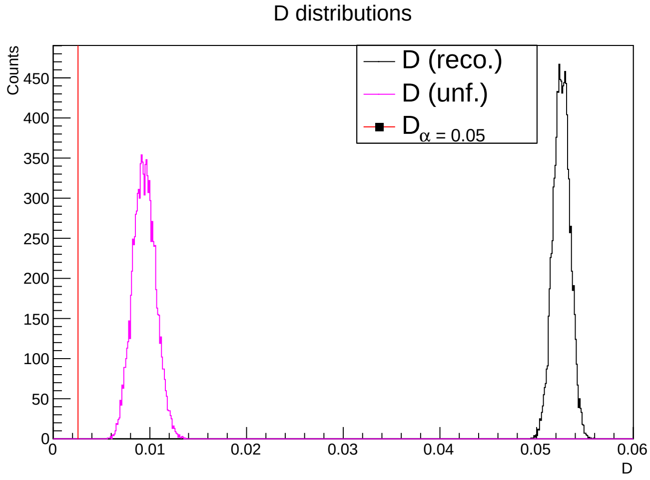

Figure 3.19: *Distributions of D when simulating* $\phi_0 = 100.0 \cdot 10^{-11}$ TeV^{-1}cm^{-2}s^{-1} *and* $\Gamma = 3$. *An improvement may be seen by the shift of the unfolded distribution (pink) towards* 0 *in regard to the reconstructed distribution (black).*

agreement with the true one. In figure 3.20 one may see the distributions of D for both the reconstructed and unfolded count spectrum. The distributions are equal within the errors, taking the full width at half maximum as errors.

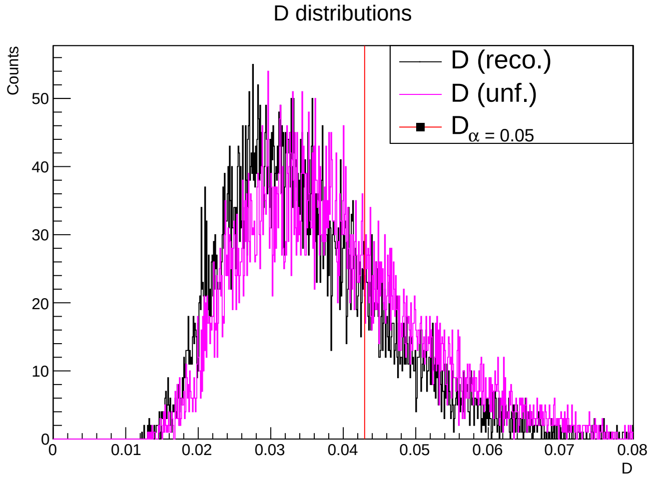

Figure 3.20: *Distributions of D when simulating $\phi_0 = 0.529$ and $\Gamma = 2$. Both distributions for the reconstructed count spectrum (black) and the unfolded count spectrum (pink) are equal within errors.*

After having tested the count spectra, the flux spectra are tested. This test is conducted by fitting function (3.11) to the reconstructed (reco.) and unfolded (unf.) flux spectra. The reconstructed and unfolded flux spectra are fitted using a log–likelihood fit assuming a Poisson count distribution within each bin. Even though the data in bins of the unfolded spectrum is not distributed following a Poisson distribution the fit is still viable, yielding the maximum likelihood for parameters which describe the data best. Additionally the reconstructed flux spectrum is fitted using a Forward Folded fit as described in subsection 3.2.1 to allow for a comparison between a Forward Folded fit and a fit to unfolded data. The fit values are the values of the fitted function at $E_0 = 1\,\mathrm{TeV}$. The best fit values are derived by filling a 2D histogram with the fit values obtained in each simulation and finding the bin with the maximum amount of counts. If more than one bin has the maximum amount of counts the one closest to the true value is chosen as best fit parameter. The errors are the binwidths of this 2D histogram. Table 3.9 shows the best fit values for ϕ_0 when fitting over the full range and safe range of the flux spectra. The safe range is defined as range in which all runs have contributed, excluding possible fluctuations at very low and very high energies.

The normalisation is recovered with best precision if a log–likelihood fit to the unfolded data is applied in all cases. The Forward Folded fit seems to not recover the normalisation well. A log–likelihood fit to the reconstructed data recovers the normalisation better in comparison to the Forward Folded fit.

Table 3.10 shows the best fit values for Γ when fitting over the full range and safe range of the flux spectra. The values and errors are obtained in the same way as mentioned earlier.

The original index is recovered well by both the Forward Folded fit to the reconstructed data as well as the log–likelihood fit to the unfolded data. The log–likelihood fit to the reconstructed data recovers the spectral index with inferior accuracy.

Except for the case of simulating low statistics and fitting the reconstructed data with a log–likelihood fit, the best fit values when fitting over safe energy range and full energy

$\phi_{0,true}$ [TeV^{-1}cm^{-2}s^{-1}]	Γ_{true}	log–likelihood (rec.)		Forward Folded (rec.)		log–likelihood (unf.)	
		full	safe	full	safe	full	safe
$100.0 \cdot 10^{-11}$	2	$98.15_{\pm0.10}$	$98.15_{\pm0.10}$	$97.05_{\pm0.10}$	$97.15_{\pm0.10}$	$99.95_{\pm0.10}$	$99.95_{\pm0.10}$
$100.0 \cdot 10^{-11}$	3	$102.45_{\pm0.10}$	$102.45_{\pm0.10}$	$93.95_{\pm0.10}$	$93.95_{\pm0.10}$	$99.65_{\pm0.10}$	$99.65_{\pm0.10}$
$0.529 \cdot 10^{-11}$	2	$0.513_{\pm0.005}$	$0.523_{\pm0.005}$	$0.518_{\pm0.005}$	$0.518_{\pm0.005}$	$0.528_{\pm0.005}$	$0.538_{\pm0.005}$
$0.356 \cdot 10^{-11}$	3	$0.358_{\pm0.005}$	$0.368_{\pm0.005}$	$0.338_{\pm0.005}$	$0.333_{\pm0.005}$	$0.358_{\pm0.005}$	$0.353_{\pm0.005}$

Table 3.9: *Best fit values of* $\phi_{0,fit}$ *in units of* $10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ *when fitting reconstructed and unfolded data over the full energy range and the safe energy range for simulations based on H.E.S.S. phase 1 data.*

$\phi_{0,true}$ [TeV^{-1}cm^{-2}s^{-1}]	Γ_{true}	log–likelihood (reco.)		Forward Folded (reco.)		log–likelihood (unf.)	
		full	safe	full	safe	full	safe
$100.0 \cdot 10^{-11}$	2	$1.983_{\pm0.002}$	$1.983_{\pm0.002}$	$2.001_{\pm0.002}$	$2.001_{\pm0.002}$	$2.001_{\pm0.002}$	$1.999_{\pm0.002}$
$100.0 \cdot 10^{-11}$	3	$2.935_{\pm0.002}$	$2.935_{\pm0.002}$	$3.003_{\pm0.002}$	$3.003_{\pm0.002}$	$3.001_{\pm0.002}$	$3.001_{\pm0.002}$
$0.529 \cdot 10^{-11}$	2	$1.994_{\pm0.004}$	$1.970_{\pm0.004}$	$1.990_{\pm0.004}$	$1.986_{\pm0.004}$	$2.002_{\pm0.004}$	$2.006_{\pm0.004}$
$0.356 \cdot 10^{-11}$	3	$2.962_{\pm0.004}$	$2.914_{\pm0.004}$	$3.006_{\pm0.004}$	$3.006_{\pm0.004}$	$2.998_{\pm0.004}$	$3.006_{\pm0.004}$

Table 3.10: *Best fit values of* Γ_{fit} *when fitting reconstructed and unfolded data over the full energy range and the safe energy range for simulations based on H.E.S.S. phase 1 data.*

range are in agreement. When simulating low statistics and fitting the reconstructed data with a log–likelihood fit, the results provided by a fit over the full energy range are superior to the ones obtained from a fit over the safe energy range. The good agreement of safe range fit values and full range fit values in the other cases may be explained by the fit algorithms taking the count spectra rather than the flux spectra into account. The count spectra have minimal values in bins outside the safe range and therefore do not heavily influence the fit.

A visual representation of the best fit values when fitting over the full energy range and their 68 %, 95 % and 99 % error intervals may be seen in figure 3.21.
The error intervals are obtained from simulations, hence their not smooth shape. When simulating an index of $\Gamma = 3$ the fit parameter seem correlated. This correlation may be due to analysing the fit parameters at $E_0 = 1\,\mathrm{TeV}$ as opposed to the point of least correlation, the Pivot energy. By definition the parameters at the Pivot energy only vary within their statistical errors, but do not vary dependent on each other. If the parameters are read out at an energy different from the Pivot energy, their values might change depending on the other parameter in order to describe the fitted flux points. The Pivot energy in this case supposedly is smaller than the readout energy E_0. If the Pivot energy is smaller than the readout energy, the normalisation at readout energy is negatively correlated to the spectral index. This correlation may be described schematically as follows. The spectrum of a source is measured twice with identical observation parameters, yielding approximately the same amount of counts in both measurements. The first measurement's flux spectrum is fitted and the parameters ϕ_1 and Γ_1 are obtained. During the second measurement statistical fluctuations cause an overshoot of counts at low energies and a

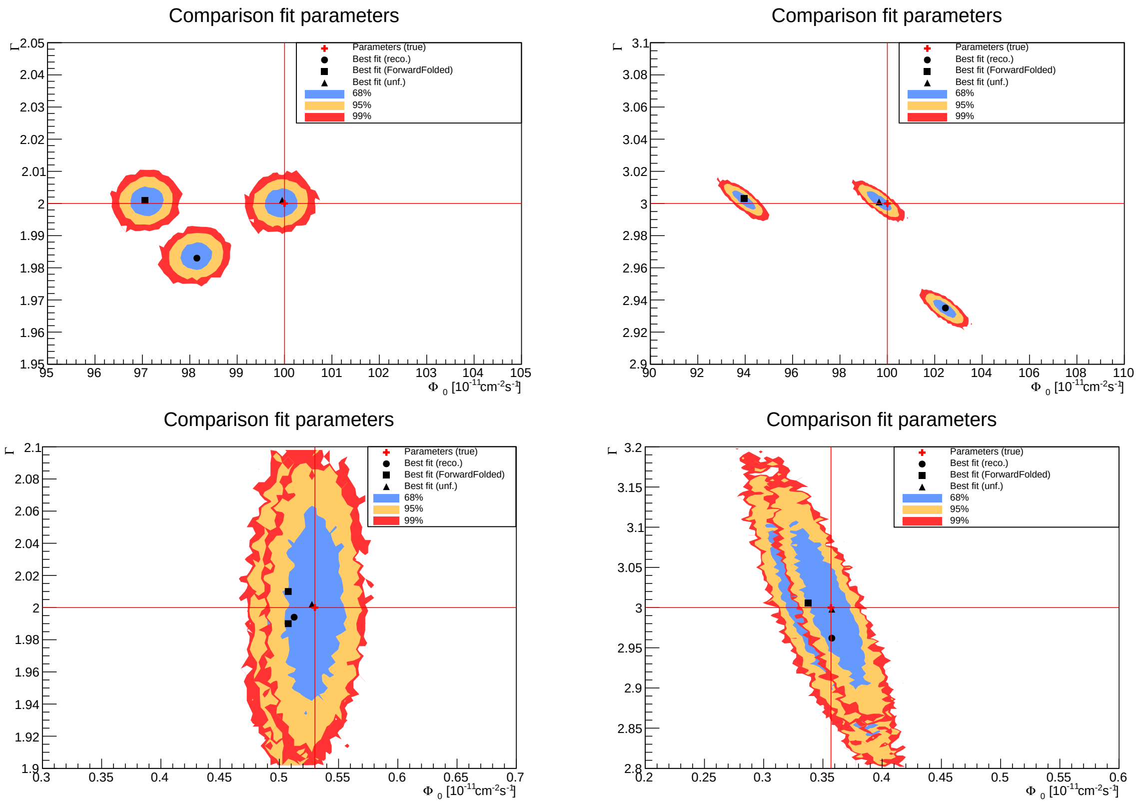

Figure 3.21: *Best fit values for different fit algortihms fitting over the entire energy range with 68 %, 95 % and 99 % error intervals, which are obtained from simulations. The red lines indicate the position of the original input parameters. The simulated spectra are based on H.E.S.S. phase 1 data. In the case of* $\phi_0 = 0.529 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ *and* $\Gamma = 2$ *two best fit values may be seen for the Forward Folded fit. The lower one is preferred due to its closer proximity to the true values.*

lack of counts at high energies, preserving the total amount of counts. The flux spectrum of the second measurement is fitted, yielding parameters ϕ_2 and Γ_2. If the parameters are read out at the Pivot energy, $\phi_1 = \phi_2$ and $\Gamma_1 = \Gamma_2$ hold true. If the parameters are read out at an energy greater than the Pivot energy, $\phi_1 > \phi_2$ and in order to describe the flux spectrum correctly $\Gamma_1 < \Gamma_2$ hold true, taking the overshoot of counts at low energies and lack of counts at high energies into account.

When choosing $\Gamma = 2$ no correlation is found. This lack of correlation may be explained by $E_0 = 1\,\mathrm{TeV}$ being close to the Pivot energy in this case. In figure 3.21 one may see that the best fit parameters when simulating realistic statistics are in agreement with the true values. The best fit values when fitting the unfolded spectrum recover the orginial parameters with best precision in all cases. The Forward Folded method recovers the spectral index with high precision, but fails to recover the normalisation. The log–likelihood fit to reconstructed data yields indices smaller than the original index and normalisations smaller than the original normalisation in the case of $\Gamma = 2$ and larger normalisations in the case of $\Gamma = 3$. This behaviour may be explained by the minor energy bias introduced at low energies. When simulating a soft spectrum ($\Gamma = 3$) more particles at low energies are affected by energy bias. The bias shifts true events at low energies towards higher energies in reconstructed energy, yielding an excess in these higher energy bins and hence a larger normalisation. When simulating a hard spectrum ($\Gamma = 2$) less particles at low energies are simulated and exposed to energy bias, yielding a smaller normalisation at $E_0 = 1\,\mathrm{TeV}$.

Simulations based on PKS 2155–304 measured with H.E.S.S. phase 1 with rebinning

Simulations based on the EMM and AT obtained from 28 measurements of PKS 2155–304 with H.E.S.S. phase 1 and their results when applying rebinning during unfolding are discussed in this paragraph. The rebinned EMM and AT may be seen in figure 3.22.

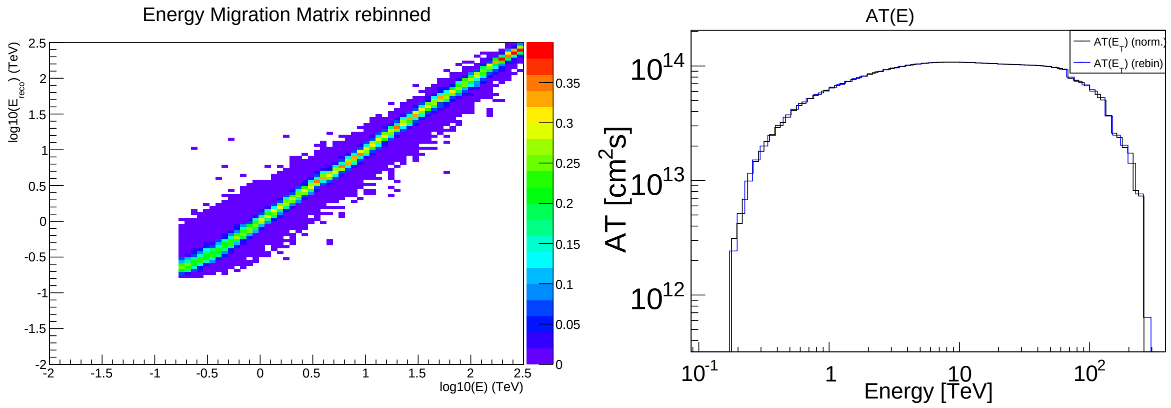

Figure 3.22: *Rebinned energy migration matrix (left) and area–time (AT) (right) derived from analysing* 28 *measurements of PKS 2155–304 with H.E.S.S. in phase 1. The binning was chosen to be* $n_T = n_R/1.4$.

The area–times are calculated from the original binning by assuming an equal distribution over each bin and assigning values according to this distribution to the new bins. This method only provides a rough estimate of the actual area–times, and hence may cause inaccuracies when examining unfolded data. A more sophisticated approach would be the filling of a rebinned area–time from lookups during analysis. This approach has not been tested during this thesis.

Since the results of K–S tests in section 3.1.2 do not allow for a conclusion whether rebinning yields comparable results, no K–S test is conducted for rebinned data–driven simulations. Instead the fit parameters obtained from fitting the rebinned unfolded spectrum are compaired. As mentioned earlier, no significant difference in fit parameters is obtained when fitting over a safe energy range in comparison to the full energy range, hence the fit parameters in the following are obtained from fitting over the full range. Table 3.11 shows the best fit parameters for ϕ_0. The best fit parameters are derived from 2D histograms as described in the previous paragraph.
The normalisation is recovered with best precision if a log–likelihood fit to the unfolded data is applied in the cases of $\phi_0 = 100 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 2$, $\phi_0 = 0.529 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 2$, and $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$, with the normalisation being equal to the one derived from a log–likelihood fit to the reconstructed data in the latter case. In the case of $\phi_0 = 100 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$ the normalisation is best recovered by a log–likelihood fit to the reconstructed

$\phi_{0,true}$ $[\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	Γ_{true}	log–likelihood (reco.)	Forward Folded (reco.)	log–likelihood (unf., rebin)
$100.0 \cdot 10^{-11}$	2	$98.15_{\pm 0.10}$	$97.05_{\pm 0.10}$	$99.25_{\pm 0.10}$
	3	$102.35_{\pm 0.10}$	$94.05_{\pm 0.10}$	$103.35_{\pm 0.10}$
$0.529 \cdot 10^{-11}$	2	$0.5175_{\pm 0.005}$	$0.5125_{\pm 0.005}$	$0.5225_{\pm 0.005}$
$0.356 \cdot 10^{-11}$	3	$0.3675_{\pm 0.005}$	$0.3325_{\pm 0.005}$	$0.3675_{\pm 0.005}$

Table 3.11: *Best fit values of* $\phi_{0,fit}$ *in units of* $10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ *when fitting reconstructed and unfolded data over the full energy range, applying rebinning during unfolding on simulations based on H.E.S.S. phase 1 data.*

data. The Forward Folded fit recovers the normalisation with the worst precision. The normalisation derived from log–likelihood fits to rebinned unfolded data yields better or comparable results in the cases of hard spectra or realistic statistics in comparison to fits to reconstructed data.

In table 3.12 the best fit values for Γ may be seen. Values and errors are derived in the same way as described earlier.

$\phi_{0,true}$ $[\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	Γ_{true}	log–likelihood (reco.)	Forward Folded (reco.)	log–likelihood (unf., rebin)
$100.0 \cdot 10^{-11}$	2	$1.983_{\pm 0.002}$	$2.001_{\pm 0.002}$	$1.973_{\pm 0.002}$
	3	$2.935_{\pm 0.002}$	$3.001_{\pm 0.002}$	$2.935_{\pm 0.002}$
$0.529 \cdot 10^{-11}$	2	$1.978_{\pm 0.004}$	$2.010_{\pm 0.004}$	$1.970_{\pm 0.004}$
$0.356 \cdot 10^{-11}$	3	$2.992_{\pm 0.004}$	$3.022_{\pm 0.004}$	$2.950_{\pm 0.004}$

Table 3.12: *Best fit values of* Γ_{fit} *when fitting reconstructed and unfolded data over the full energy range, applying rebinning during unfolding on simulations based on H.E.S.S. phase 1 data.*

The original index is recovered best by applying a Forward Folded fit to reconstructed data. Except for the case $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$ the indices derived from a log–likelihood fit to unfolded data are in least agreement to the original index.
A visual representation of the best fit values when fitting over the full energy range and their 68 %, 95 % and 99 % error intervals may be seen in figure 3.23.
In the cases of $\Gamma = 3$ the same apparent correlation as in figure 3.21 is present, which may be explained in the same way by not taking the fit parameters at the point of least correlation. The results of the log–likelihood fit and Forward Folded fit to reconstructed data perfectly align with the results obtained in previous simulations. The fit to rebinned unfolded data delivers inferior results in all cases in comparison to not rebinned data. Except for the case of $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$ fitting to rebinned unfolded data yields the value closest to the true normalisation. The spectral index Γ obtained from fits to rebinned unfolded data is inferior to the value obtained from fits to reconstructed data, except for the case $\phi_0 = 0.356 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$. In this case a log–likelihood fit to reconstructed data yields worse results. When simulating realistic statistics all fits are in agreement with the true values. In these cases a Forward Folded fit yields the best results.

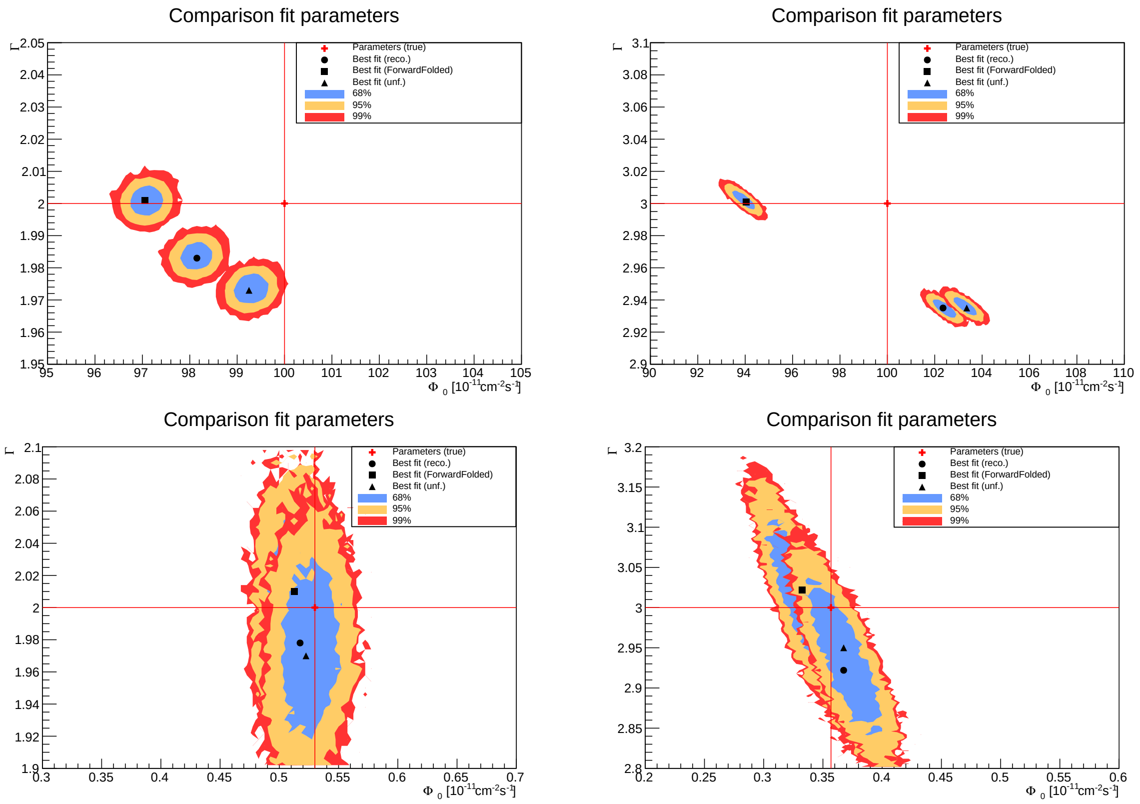

Figure 3.23: *Best fit values for different fit algortihms fitting over the entire energy range with* 68 %, 95 % *and* 99 % *error intervals, which are obtained from simulations, with rebinning during unfolding. The red lines indicate the position of the original input parameters. The simulated spectra are based on H.E.S.S. phase 1 data.*

Even though in the cases of realistic statistics the results obtained from fits to rebinned unfolded data are in agreement to the true values, the rebinned data is inferior to data with contained binning. This inferiority may be caused by the implementation of rebinning, in detail the assumption of equal distributions over bins when calculating area–times. In the current implementation rebinning during unfolding is not recommended.

Simulations based on PKS 2155–304 measured with H.E.S.S. phase 2

In this paragraph simulations based on 61 measurements of PKS 2155–304 with H.E.S.S. during phase 2 are unfolded and their results discussed. The EMM and effective area–times used in these simulations may be seen in figure 3.24.

The EMM of phase 2 covers a wider energy range than the one in phase 1, but introduces a greater bias for low and high energies. These properties allow for a harder test of the unfolding algorithm.

In table 3.13 the results of the K–S test of reconstructed and unfolded count spectra over full energy range may be seen.

In the case of $\phi_0 = 100.0 \cdot 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $\Gamma = 3$ the unfolded distribution is superior to the reconstructed one, even though both entries in the table are 0. Figure 3.25 shows both distributions of D.

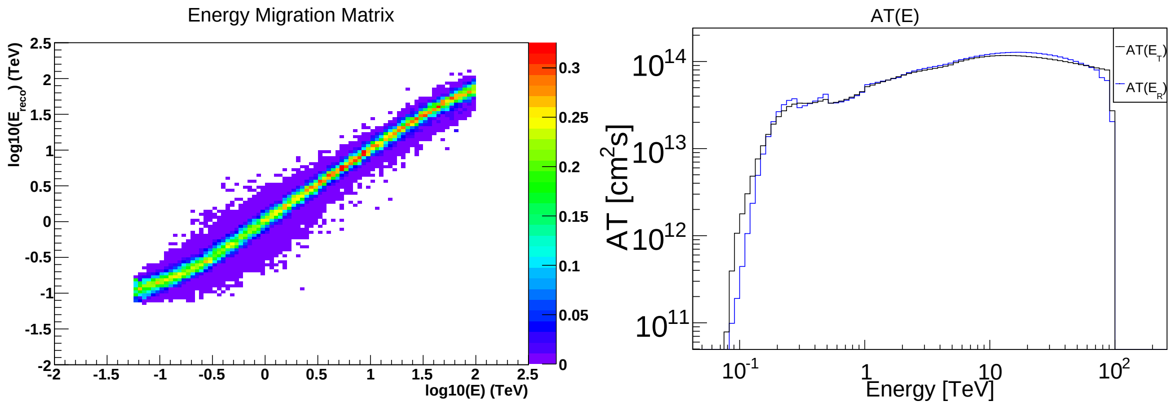

Figure 3.24: *Energy migration matrix (EMM) (left) and area–time (AT) (right) derived from analysing* 61 *measurements of PKS 2155–304 with H.E.S.S. in phase 2.*

ϕ_0 [TeV^{-1}cm^{-2}s^{-1}]	Γ	reco.	unf.
$100.0 \cdot 10^{-11}$	2	0 %	38.3 %
	3	0 %	0 %
$0.529 \cdot 10^{-11}$	2	5.9 %	65.7 %
$0.356 \cdot 10^{-11}$	3	0 %	61.6 %

Table 3.13: *Percentage of D contained within the* 5 % *significance level obtained from a K–S test on the true distribution. The reconstructed (reco.) or unfolded (unf.) count spectrum is in agreement with the true distribution if at maximum* 5 % *are not contained within the significance limit* $D_{0.05}$. *The simulated count spectra are based on H.E.S.S. phase 2 data.*

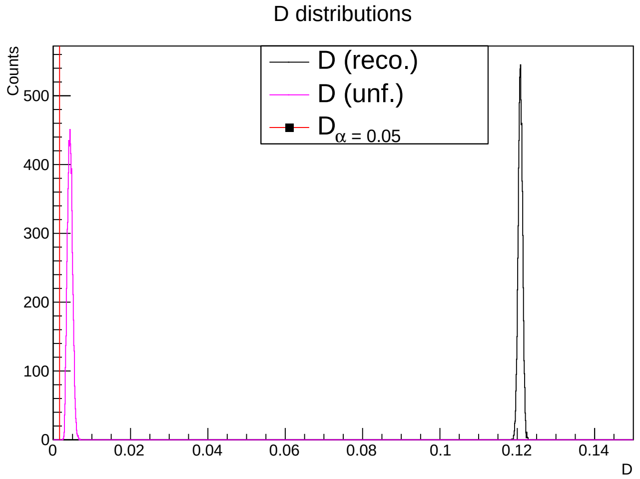

Figure 3.25: *Distributions of D when simulating* $\phi_0 = 100.0 \cdot 10^{-11}$ TeV^{-1}cm^{-2}s^{-1} *and* $\Gamma = 3$. *An improvement may be seen by the shift of the unfolded distribution (pink) towards* 0 *in regard to the reconstructed distribution (black). The simlations are based on H.E.S.S. phase 2 data.*

The data in table 3.13 may be interpreted as follows. Due to the greater bias in the EMM of H.E.S.S. phase 2 the unfolded spectra yield better results than the reconstructed

spectra for all sets of parameters.
Whether this result holds true when fitting flux spectra is tested in the following. The fits are conducted as described previously. Table 3.14 shows the best fit values for normalisations ϕ_0 when fitting over full and safe energy range. The best fit values are obtained from 2D histograms as described in the previous paragraphs.

$\phi_{0,true}$ [TeV^{-1}cm^{-2}s^{-1}]	Γ_{true}	log–likelihood (reco.)		Forward Folded (reco.)		log–likelihood (unf.)	
		full	safe	full	safe	full	safe
$100.0 \cdot 10^{-11}$	2	$99.75_{\pm0.10}$	$99.75_{\pm0.10}$	$92.65_{\pm0.10}$	$92.65_{\pm0.10}$	$99.95_{\pm0.10}$	$99.85_{\pm0.10}$
$100.0 \cdot 10^{-11}$	3	$110.85_{\pm0.10}$	$110.85_{\pm0.10}$	$89.65_{\pm0.10}$	$89.95_{\pm0.10}$	$99.65_{\pm0.10}$	$99.65_{\pm0.10}$
$0.529 \cdot 10^{-11}$	2	$0.527_{\pm0.003}$	$0.532_{\pm0.003}$	$0.489_{\pm0.003}$	$0.492_{\pm0.003}$	$0.527_{\pm0.003}$	$0.529_{\pm0.003}$
$0.356 \cdot 10^{-11}$	3	$0.398_{\pm0.003}$	$0.393_{\pm0.003}$	$0.323_{\pm0.003}$	$0.323_{\pm0.003}$	$0.358_{\pm0.003}$	$0.353_{\pm0.003}$

Table 3.14: *Best fit values of* $\phi_{0,fit}$ *in units of* 10^{-11} TeV^{-1}cm^{-2}s^{-1} *when fitting reconstructed and unfolded data over the full energy range and the safe energy range for simulations based on H.E.S.S. phase 2 data.*

A log–likelihood fit to the unfolded data yields the results in best agreement to the true values for all sets of parameters. The normalisation obtained from a Forward Folded fit to reconstructed data is in least agreement to the true normalisation except for $\phi_0 = 0.356 \cdot 10^{-11}$ TeV^{-1}cm^{-2}s^{-1} and $\Gamma = 3$. In this case the result obtained from fitting reconstructed data using a log–likelihood fit is inferior. The differences when fitting over full or safe energy range are negligible.
In table 3.15 the best fit values of Γ when fitting over full and safe range may be seen.

$\phi_{0,true}$ [TeV^{-1}cm^{-2}s^{-1}]	Γ_{true}	log–likelihood (reco.)		Forward Folded (reco.)		log–likelihood (unf.)	
		full	safe	full	safe	full	safe
$100.0 \cdot 10^{-11}$	2	$1.999_{\pm0.002}$	$1.999_{\pm0.002}$	$1.993_{\pm0.002}$	$1.993_{\pm0.002}$	$1.999_{\pm0.002}$	$2.001_{\pm0.002}$
$100.0 \cdot 10^{-11}$	3	$2.973_{\pm0.002}$	$2.973_{\pm0.002}$	$2.989_{\pm0.002}$	$2.987_{\pm0.002}$	$3.001_{\pm0.002}$	$3.001_{\pm0.002}$
$0.529 \cdot 10^{-11}$	2	$1.999_{\pm0.002}$	$1.997_{\pm0.002}$	$1.991_{\pm0.002}$	$1.987_{\pm0.002}$	$1.993_{\pm0.002}$	$1.985_{\pm0.002}$
$0.356 \cdot 10^{-11}$	3	$2.970_{\pm0.002}$	$2.982_{\pm0.002}$	$2.982_{\pm0.002}$	$2.986_{\pm0.002}$	$2.994_{\pm0.002}$	$3.014_{\pm0.002}$

Table 3.15: *Best fit values of* Γ_{fit} *when fitting reconstructed and unfolded data over the full energy range and the safe energy range for simulations based on H.E.S.S. phase 2 data.*

The data in table 3.15 allows for following conclusions. The log–likelihood fit to unfolded data yields the values closest to the true Γ in most cases. The exception is $\phi_0 = 0.529 \cdot 10^{-11}$ TeV^{-1}cm^{-2}s^{-1} and $\Gamma = 2$. In this case the spectrum is harder and less counts are affected by the high energy bias at low energies. The log–likelihood fit to reconstructed data yields a better result in this case. When simulating a spectral index of $\Gamma = 2$ log–likelihood fits to reconstructed data yield better results than Forward Folded fits. When simulating soft spectra with $\Gamma = 3$ Forward Folded fits yield better results than log–likelihood fits to reconstructed data. Fitting over the safe energy range yields comparable or worse results for Γ in most cases. The exception is $\phi_0 = 0.356 \cdot 10^{-11}$ TeV^{-1}cm^{-2}s^{-1} and $\Gamma = 3$ when fitting reconstructed data with either a log–likelihood or Forward Folded fit. Even in these cases the difference in values is small, hence a fit over the full energy range is preferred.

Figure 3.26 shows the best fit values for all sets of parameters when fitting over the full energy range and their 68 %, 95 % and 99 % error intervals.

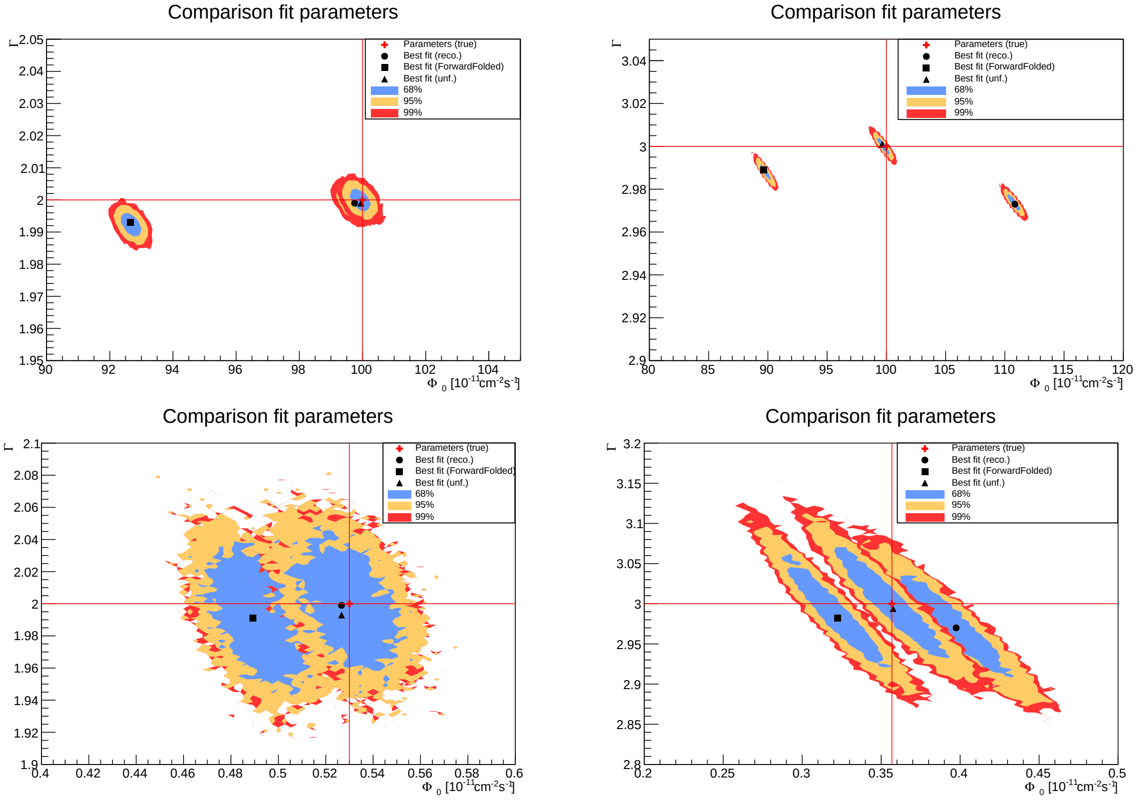

Figure 3.26: *Best fit values for different fit algortihms fitting over the entire energy range with 68 %, 95 % and 99 % error intervals, which are obtained from simulations. The red lines indicate the position of the original input parameters. The simulated spectra are based on H.E.S.S. phase 2 data.*

The apparent correlation when fitting spectra with $\Gamma = 3$ may be explained by the evaluation energy $E_0 = 1\,\text{TeV}$, which is not the energy of least correlation. The conclusions drawn from these plots are the following. In the case of hard spectra and small spectral indices ($\Gamma = 2$) the log–likelihood fits to reconstructed and unfolded data are in agreement. This agreement may be explained by the small influence energy bias has on hard spectra. In the case of soft spectra the log–likelihood fits to unfolded data yield the best results. Forward Folded fits to reconstructed data are not preferred in any case.

The overall results of fitting data–driven simulations using H.E.S.S. phase 2 data as input may be summarised as follows. In most cases the best results are achieved if a log–likelihood fit to unfolded data is conducted. The sole exception is when simulating a hard spectrum with realistic statistics when fitting the reconstructed data with a log–likelihood fit yields slightly better results. In this case a hard spectrum with realistic statistics is generated. Due to the hardness of the spectrum the majority of events is not influenced by the non–negligible energy bias at low energies, yielding a well reconstructed spectrum. In this case the fit results of log–likelihood fits to reconstructed and unfolded data do not differ significantly. When simulating soft spectra, which are highly affected by energy bias, unfolding yields superior results.

Real data from astrophysical sources with both hard and soft spectral indices is subject to test in chapter 4.

4. Unfolding of astrophysical source spectra

In this chapter unfolding of real astrophysical source spectra is conducted. In detail the algorithm is applied to data of PKS 2155–304 and the Crab nebula measured with H.E.S.S. in phase 1 and to data of PKS 2155–304 measured with H.E.S.S. in phase 2.
Additionally the stability of the unfolding algorithm in the presence of background is tested. The test is conducted by simulating measurements of the sources named above in the following way. Each bin in N_R^{ON} and N_R^{OFF} is assigned a random value within the Poisson errors of the respective bin. The resulting ON and OFF count spectra are unfolded separately and subsequently combined to obtain flux spectra. From each set of data 1000 measurements are simulated and unfolded with $n_H = 500$.
Some runs of PKS 2155–304 and the Crab nebula, which are listed as phase 1 runs in this work, are measured after the addition of the fifth telescope to H.E.S.S. and could be classified as phase 2 runs. The runs analysed are measured without taking data of the fifth telescope into account, yielding an instrument response in agreement to the one obtained during H.E.S.S. phase 1. For simplicity reasons these runs are denoted as phase 1 runs in this work.

4.1 Crab nebula measured with H.E.S.S. in phase 1

In this section the unfolding algorithm is applied to the combined data of 27 runs of the Crab nebula with H.E.S.S. in phase 1 measured from October 2010 to March 2013 and to simulations based on this data. The runs are selected with spectral selection cuts and analysed with the standard analysis configuration of hap. The unfolded and reconstructed flux spectra are fitted with a simple power–law and compared to literature values (Aharonian et al., 2008).

In a first step the algorithm is applied to real data. The unfolding und subsequent fitting of real data is carried out 1000 times to estimate the statistical error induced by the unfolding algorithm. In table 4.1 the best fit results when applying log–likelihood fits to reconstructed (reco.) and unfolded (unf.) spectra and Forward Folded fits to reconstructed data are shown. The best fit values are obtained from filling a 2D histogram, seen in figure 4.2, and finding the bin with the maximum amount of counts. The errors are obtained from the fit.
The best fit values of all three fits are in agreement within their errors.
After the unfolding of real data the algorithm is applied to simulated measurements. As described previously 1000 simulations are carried out, assigning random values of to each bin in the count spectra. These count spectra are unfolded simulating $n_H = 500$ mock

fit	ϕ_0 $[10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	Γ
log–likelihood (reco.)	3.2495 ± 0.0686	2.6075 ± 0.0229
Forward Folded (reco.)	3.2965 ± 0.0686	2.6595 ± 0.0240
log–likelihood (unf.)	3.3755 ± 0.0704	2.6465 ± 0.0230
literature	$3.45 \pm 0.05_{Stat}$	$2.63 \pm 0.02_{Stat}$

Table 4.1: *Best fit values when applying various fit algorithms to the combined reconstructed spectra of* 27 *runs measuring the Crab nebula with H.E.S.S. in phase 1 and a log–likelihood fit to their unfolded counterparts and literature values (Aharonian et al., 2008). The errors are obtained from the fit.*

spectra. From the simulated count spectra flux spectra are derived, which subsequently are fitted. By definition the expected mean of the simulations is the real data. In figure 4.1 the best fit values from table 4.1 and the 68 %, 95 % and 99 % error intervals for the log–likelihood fit to unfolded data obtained from simulations may be seen. The non uniform shape of the error intervals is due to the limited number of simulations.

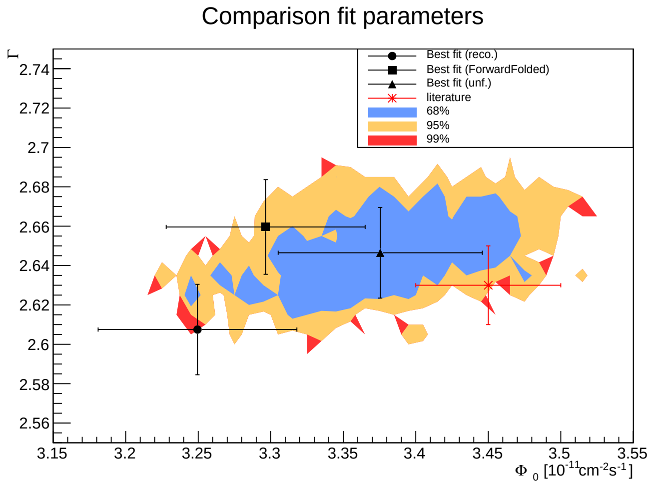

Figure 4.1: *Fit values for different fit algortihms fitting over the entire energy range with error bars obtained from the fit. The red lines indicate the position of the literature values (Aharonian et al., 2008). The* 68 %, 95 % *and* 99 % *error intervals for the log–likelihood fit to unfolded data are obtained from simulating* 1000 *measurements based on real data and subsequent fitting.*

The error bars obtained from the fit and the error intervals from simulations are in agreement. The shape of error intervals suggests no correlation of the fitting parameters. This lack of correlation may be explained by reading out the fit parameters at their Pivot energy $E_{Pivot} = E_0 = 1\,\mathrm{TeV}$. The log–likelihood fit to the unfolded spectrum is in best agreement with the literature values. Considering the simple analysis, large errors on the fits and different fit ranges of [0.41 TeV; 40.0 TeV] in the literature and [0.01 TeV; 300.0 TeV] in this work, only taking bins with entries larger 0 into account, a fair result is obtained when unfolding the measured data. A more refined analysis is supposed to yield even better results.

Figure 4.2 shows the 68 %, 95 % and 99 % error intervals for the log–likelihood fit to unfolded data when unfolding the same set of measured data 1000 times. Additionally literature values (Aharonian et al., 2008) and best fit parameters obtained from a simple power–law fit to reconstructed and unfolded flux spectra using log–likelihood fits to reconstructed (reco.) and unfolded (unf.) spectra and Forward Folded fits to the reconstructed spectra are displayed.

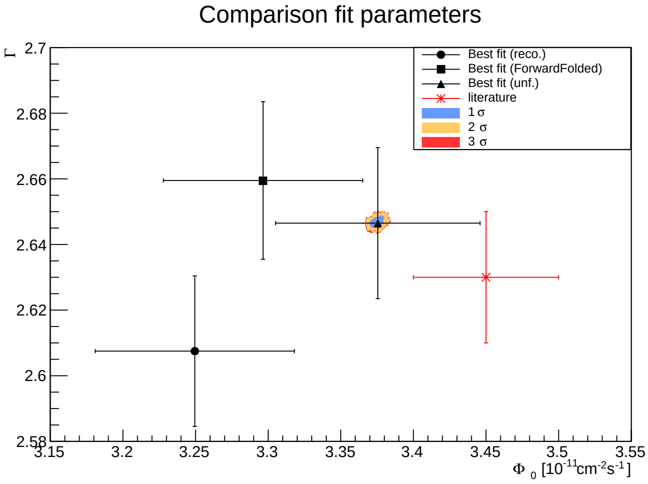

Figure 4.2: *Fit values for different fit algortihms fitting over the entire energy range with errors obtained from the fit. The red lines indicate the position of the literature values (Aharonian et al., 2008). The* 68 % *,* 95 % *and* 99 % *error intervals for the fit to unfolded data are obtained from unfolding the measured data* 1000 *times and subsequent fitting.*

The error intervals derived from applying the unfolding algorithm to the same set of data 1000 are well contained within the errors obtained from the fit. This result proves the stability of the unfolding algorithm. Applying the unfolding algorithm only once to a set of data does not introduce significant errors and is hence viable in all cases.

4.2 PKS 2155–304 measured with H.E.S.S. in phase 1

In this section data obtained from 28 runs measuring PKS 2155–304 with H.E.S.S. in phase 1 from October 2010 to November 2012 is unfolded and used for simulating measurements. The runs are selected with spectral selection cuts and the data is analysed with the standard analysis configuration of hap. The flux spectra obtained from measurement and unfolding are fitted, using log–likelihood and Forward Folded fits. The fit function is a simple power–law in all cases.

In order to obtain best fit values real data is unfolded 1000 times and the fit parameters are filled in a 2D histogram seen in figure 4.4. The best fit values are obtained from finding the bin with the maximum amount of counts.
Table 4.2 shows the best fit values when applying log–likelihood fits to reconstructed (reco.) and unfolded (unf.) flux spectra and when applying a Forward Folded fit to the reconstructed flux spectrum. Additionally the literature values are displayed. The literature values were obtained applying a simple power–law fit to reconstructed data of PKS 2155–304 measured in 2003 (Aharonian et al., 2005).
The best fit values of all three fits are in agreement within their errors.

fit	ϕ_0 $[10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	Γ
log–likelihood (reco.)	0.1845 ± 0.0161	3.202 ± 0.0894
Forward Folded (reco.)	0.1615 ± 0.0165	3.282 ± 0.1014
log–likelihood (unf.)	0.1735 ± 0.0161	3.278 ± 0.0931
literature	0.273 ± 0.017	3.37 ± 0.17

Table 4.2: *Best fit values when applying various fit algorithms to the combined reconstructed spectra of* 28 *runs measuring PKS 2155–304 with H.E.S.S. in phase 1 and a log–likelihood fit to their unfolded counterparts and literature values (Aharonian et al., 2005). The errors are obtained from the fit.*

The flux normalisation ϕ_0 when fitting the reconstructed spectrum with a log–likelihood fit deviates least from the literature value. This result aligns with the results obtained from simulations in subsection 3.2.3. The spectral index of a Forward Folded fit to the reconstructed spectrum is in best agreement with the literature value, yet the result is comparable to the one obtained from fitting the unfolded spectrum with a log–likelihood fit. The log–likelihood fit to unfolded data yields less deviation from the literature normalisation than a Forward Folded fit and better agreement to the literature spectral index than a log–likelihood fit to reconstructed data. The fit results of the respective other parameter are less in agreement to literature in both cases.

The stability of the unfolding algorithm is tested by simulating measurements as described previously. Figure 4.3 shows the best fit values from table 4.2, the 68 %, 95 % and 99 % error intervals for the log–likelihood fit to unfolded data obtained from simulations and literature values (Aharonian et al., 2005).

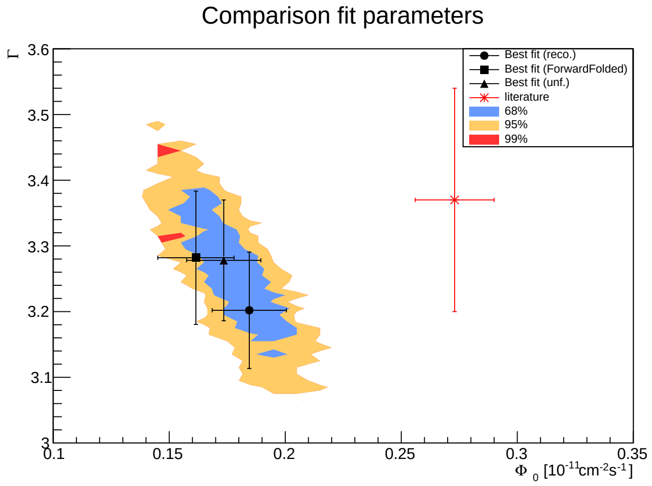

Figure 4.3: *Fit values for different fit algortihms fitting over the entire energy range with errors obtained from the fit. The red lines indicate the position of the literature values measured in 2003 (Aharonian et al., 2005). The* 68 %, 95 % *and* 99 % *error intervals are obtained from simulating* 1000 *measurements based on real data with subsequent unfolding and fitting.*

The non uniform shape of the error intervals may be explained by the limited amount of simulations. The errors obtained from the fit and the error intervals from simulations

align well. The error intervals suggest a correlation of ϕ_0 and Γ in the sense of increasing values of ϕ_0 correlating with decreasing values of Γ. The reason for this correlation is choosing a readout energy of $E_0 = 1\,\mathrm{TeV}$ and not the Pivot energy $E_{Pivot} = 0.4429\,\mathrm{TeV}$. The negative correlation may be explained in the same way as described in subsection 3.2.3.
The variability of the source complicates the choice of appropriate literature values. This choice may be the main reason for disagreement of values derived in this work and literature values. The fitted energy range may also play a significant role. The literature values were obtained from fitting in the range $[0.2\,\mathrm{TeV}; 10.5\,\mathrm{TeV}]$ while the best fit values in this thesis are obtained from fitting over an energy range of $[0.01\,\mathrm{TeV}; 300.0\,\mathrm{TeV}]$ only taking bins with entries larger 0 into account.

In figure 4.4 the best fit values as well as the 68 %, 95 % and 99 % error intervals for the log–likelihood fit to unfolded data and literature values (Aharonian et al., 2005) are shown. The error intervals are obtained from unfolding the same set of measured data 1000 times and fitting it subsequently.

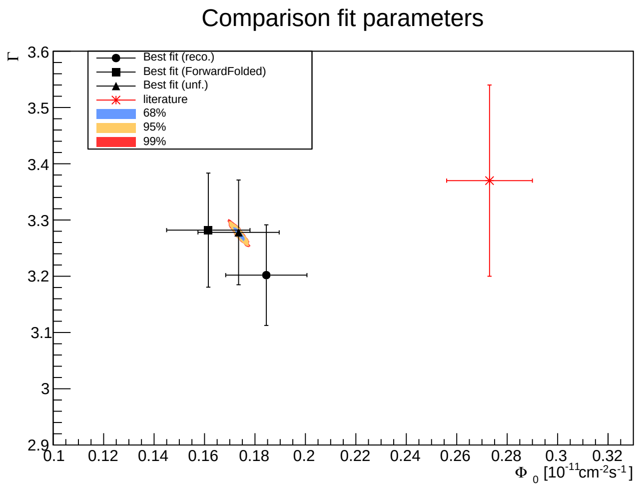

Figure 4.4: *Fit values for different fit algortihms fitting over the entire energy range with errors obtained from the fit. The red lines indicate the position of the literature values measured in 2003 (Aharonian et al., 2005). The* 68 %, 95 % *and* 99 % *error intervals are obtained unfolding measured data* 1000 *times and subsequent fitting.*

The error intervals are well contained within the errors obtained from the fit. As described in the previous section unfolding a set of measured data only once is viable.

4.3 PKS 2155-304 measured with H.E.S.S. in phase 2

In this section unfolding is applied to data from a combined measurement of 61 runs of PKS 2155–304 measured with H.E.S.S. in phase 2 in between May 2014 and July 2015 and to simulations based on this data. The runs are selected with spectral selection cuts and analysed with the standard analysis configuration of hap. The unfolded and reconstructed flux spectra are fitted with a simple power–law and compared to literature values (HESS Collaboration, 2016).

The literature values were obtained from analysing data from 2013 to 2014 only using monoscopic events. Monoscopic events are events only recorded by the fifth, largest telescope of H.E.S.S., which has been added to the array in 2012. The addition of this last telescope is the cause of the change from phase 1 to phase 2. The data analysis in this work is making use of all five telescopes of H.E.S.S.

Table 4.3 shows the best fit results of log–likelihood fits to reconstructed (reco.) and unfolded (unf.) flux spectra and Forward Folded fits to the reconstructed spectrum. The fit values are obtained from unfolding the measured data 1000 times and fitting the unfolded flux spectra subsequently. The best fit values are determined by filling all fit values in a 2D histogram, seen in figure 4.6, and finding the bin with the maximum amount of counts. The literature value has been obtained from fitting a log–parabola function $dN/dE = \phi_0(E/E_0)^{-\Gamma-\beta\cdot\log(E/E_0)}$ to monoscopic data (HESS Collaboration, 2016). The literature values are given at the decorrelation energy $E = 156\,\mathrm{GeV}$.

fit	ϕ_0 [$10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	Γ
log–likelihood (reco.)	0.4315 ± 0.0220	2.830 ± 0.0366
Forward Folded (reco.)	0.3655 ± 0.0203	2.818 ± 0.0368
log–likelihood (unf.)	0.4405 ± 0.0212	2.670 ± 0.0320
literature	0.511 ± 0.15	2.63 ± 0.07

Table 4.3: *Best fit values when applying various fit algorithms to the combined reconstructed spectra of* 61 *runs measuring PKS 2155–304 with H.E.S.S. in phase 2 and a log–likelihood fit to their unfolded counterparts and literature values (HESS Collaboration, 2016) obtained from a log–parabola fit to monoscopic data. The errors are obtained from the fit.*

The best fit values are not in agreement within their errors in all cases. This result aligns with the results obtained from simulations of large statistics discussed in subsection 3.2.3. A comparison to simulations with large statistics is viable in this case since a large number of runs is combined to obtain the data set examined in this section.

Even though the methods used in HESS Collaboration (2016) differ heavily from the ones used in this work, the log–likelihood fit to unfolded data is in fair agreement to the literature values. The difference in normalisation may be explained by the literature value being readout at the decorrelation energy rather than at $E_0 = 1\,\mathrm{TeV}$ and being fitted over an energy range of $[0.08\,\mathrm{TeV}; 1.2\,\mathrm{TeV}]$ rather than $[0.01\,\mathrm{TeV}; 300.0\,\mathrm{TeV}]$ only taking bins with entries larger 0 into account as used in this work. Due to large discrepancies in analysis methods the agreement of unfolded and literature values may prove to be wrong.

The stability of the unfolding algorithm in the presence of background is tested by simulating measurements as described previously. In figure 4.5 the best fit values from table 4.3 and the 68 %, 95 % and 99 % error intervals for the log–likelihood fit to unfolded data obtained from simulations are shown. Additionally the literature values are displayed.

The non uniform shape of the error intervals is caused by the limited amount of simulations. The error intervals are well contained within the error bars obtained from the fit. The shape of the error intervals suggests ϕ_0 to be negatively correlated to Γ, yielding smaller values of Γ for higher values of ϕ_0, due to reading out the parameters at $E_0 = 1\,\mathrm{TeV}$ rather than $E_{Pivot} = 0.3321\,\mathrm{TeV}$. The best fit value obtained from fitting a log–likelihood fit to unfolded data yields the best agreement with the literature values. This agreement

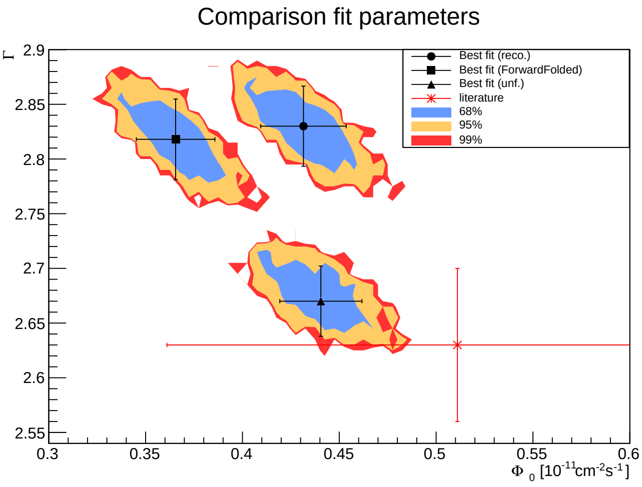

Figure 4.5: *Fit values for different fit algortihms fitting over the entire energy range with error bars obtained from the fit. The red lines indicate the position of the literature values (HESS Collaboration, 2016). The* 68 % *,* 95 % *and* 99 % *error intervals are obtained from simulating* 1000 *measurements and subsequent fitting.*

of unfolded data and literature data may prove to be wrong. A more legitimate comparison may be achieved by fitting the data in this work using a log–parabola function. The fit parameters obtained from fitting the reconstructed spectrum with a log–likelihood and Forward Folded fit are in agreement with expectation from simulations in subsection 3.2.3.

Figure 4.6 shows the best fit values obtained from unfolding and fitting the same set of measured data 1000 times. The error intervals are well contained within the error bars obtained from the fit. As described in previous sections unfolding a measured data set only once yields viable results.

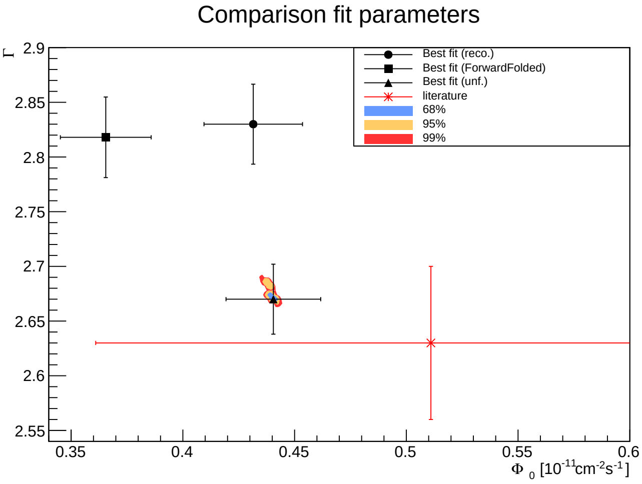

Figure 4.6: *Fit values for different fit algortihms fitting over the entire energy range with errors obtained from the fit. The red lines indicate the position of the literature values (HESS Collaboration, 2016). The* 68 % *,* 95 % *and* 99 % *error intervals are obtained from unfolding the same set of measrued data* 1000 *times.*

5. Conclusion

In this thesis a Bayesian Unfolding algorithm has been tested and implemented into the Heidelberg Analysis Program (hap).

The algorithm has been tested for simple block, Gaussian and halfcircle distributions. The simulated reconstructed and unfolded distributions have been validated using a Kolmogorov-Smirnov test in order to determine the quality of unfolding. In all cases unfolding failed to fully reconstruct the original distributions, but still yielded an improvement in comparison to reconstructed distributions. The resolution of the instrument response was found to heavily influence both the number of iterations until abortion as well as the accuracy of the unfolded distribution, requiring an increasing amount of iterations until the abort critertion is met and yielding a decreasing accuracy with deteriorating resolution. The algorithm was found to be stable in the presence of bias. The amount of simulated statistics was not found to have a significant impact on the algorithm. Applying rebinning during unfolding was not found to be viable in all cases, especially for large statistics.
In a next step the unfolding algorithm has been tested on data–like simulations. When performing data–like simulations the unfolding algorithm failed to fully reconstruct the simulated true count spectrum, but still yielded an improvement in comparison to the reconstructed count spectrum. A conversion of simulated reconstructed and unfolded count spectra into flux spectra allowed for a comparison of different fitting algorithms, in particular for a comparison of a log–likelihood fit to unfolded data and a Forward Folded fit to reconstructed data. The log–likelihood fit to unfolded data proved to be the superior choice in all tested cases.
Rebinning during unfolding has also been tested with data–like simulations, but was found to be not desirable. An improved implementation of rebinning may change this result.

The algorithm has also been applied to real data, namely from PKS 2155–304 and the Crab nebula, and compared to literature values. Data from PKS 2155–304 has been unfolded for both H.E.S.S. phase 1 and 2. In all cases a log–likelihood fit to unfolded data yielded the best agreement with literature values compared to log–likelihood and Forward Folded fits to reconstructed data. Additionally the unfolding algorithm was found to be stable in the sense that unfolding one set of data only once yields viable results.
Further refinement of data analysis may provide even superior results.

The main goal of enabling the analysis of highly biased data measured with H.E.S.S. in phase 2 within the Heidelberg Analysis Program has been reached. The algorithm may be used for data analysis of sources with unknown spectral shape or for the analysis of sources with known spectral shape in order to obtain possibly more precise results than the existing methods provide.

Appendices

A. User manual for applying Bayesian Unfolding in hap

The method of unfolding is supposed to make a wide energy range accessible for analysis. In order to analyse data in the widest range possible, a flag *UnfoldingDesired* should be set in the configuration file of the analysis. The flag is set under [Spectrum] options as may be seen in the following.

UnfoldingDesired = true

Setting this flag is equivalent to setting the allowed energy bias to 200 % as opposed to the usual 10 %. A safe energy range is determined by deriving the minimum range in which all spectra contribute to the combined spectrum.

The application of the Bayesian Unfolding algorithm is implemented in the FitSpectrum binary of hap, which provides a convenient way of fitting energy flux spectra. The unfolded data may be fitted directly by using the following options.

Option	data type	application
FitUnfoldedData	bool	A fit to unfolded data is conducted if *FitUnfoldedData* is set to *true*. The fit algorithm should be changed to *FineBinLogLike*, the log–likelihood fit algorithm used in this thesis. The spectrum is unfolded until the abort criterion is met or the upper limit of 20 iterations is reached. The unfolding is conducted simulating 500 mock spectra.
RebinDuringUnfolding	bool	The spectrum is rebinned during unfolding if *RebinDuringUnfolding* is set to *true*. A rebinned energy migration matrix and an estimated rebinned area–time is used. *RebinDuringUnfolding* only works in combination with *FitUnfoldedData*. In the current implementation an application of rebinning during unfolding is not recommended.

An example of a fit to unfolded data from 28 runs of PKS 2155–304 with H.E.S.S. in phase 1 without rebinning during unfolding may be seen in figure A.1.
Another option is only plotting the unfolded spectrum while fitting the reconstructed spectrum, e.g. using Forward Folding. This option is described in the following table.

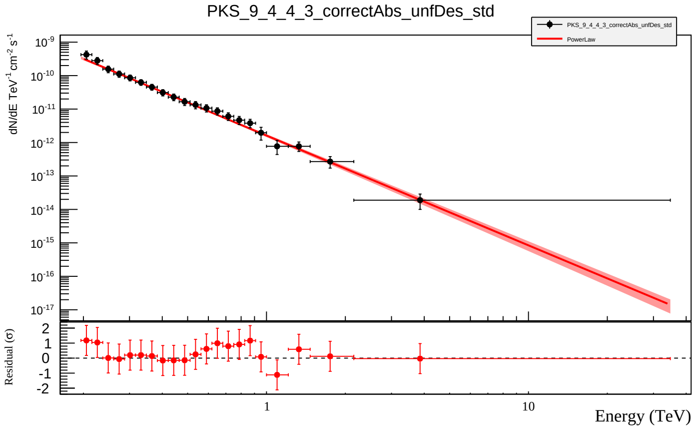

Figure A.1: *A fitted unfolded spectrum of* 28 *runs measuring PKS 2155–304 with H.E.S.S. in phase 1 using the* ***FitUnfoldedData*** *option and its residuals. The fit algorithm is* ***FineBinLogLike****. The fit is conducted over the full energy range. The errors are obtained from unfolding and hence correlated.*

Option	data type	application
PlotOption	String	Adding U to the plot options (standard: RL) will plot the unfolded data, unfolding up to a maximum of 20 iterations simulating $n_H = 500$ mock spectra. Do not combine with *FitUnfoldedData*. The reconstructed flux points are fitted in this case, but only the unfolded flux points are plotted.

An example of a Forward Folded fit to reconstructed data from 28 runs of PKS 2155–304 with H.E.S.S. in phase 1 while plotting the unfolded spectrum may be seen in figure A.2.

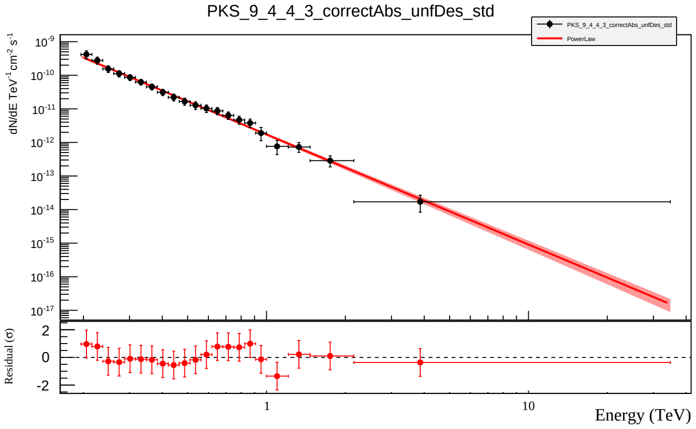

Figure A.2: *A fitted spectrum of* 28 *runs measuring PKS 2155–304 with H.E.S.S. in phase 1 using the PlotOption* ***RLU****. The fit algorithm is* ***ForwardFolded****. The fit is conducted over the full energy range. The errors are obtained from unfolding and hence correlated.*

In all cases the statistics plots of the original, reconstructed spectrum are plotted.

Bibliography

Aharonian, F. et al. (2005), 'Mulit–wavelength observations of pks 2155–304 with hess', *A&A* **442**, 895–907.

Aharonian, F. et al. (2008), 'Observations of the crab nebula with h.e.s.s.', *A&A* **457**, 899–915.

Albert, J. et al. (2007), 'Unfolding of differential energy spectra in the magic experiment', *Nuclear instruments and methods in physics research* **583**, 494ff.

D. Wouters, P. B. (n.d.), Unfolding methods for spectral reconstruction with hess. H.E.S.S. internal note.

Funk, S. (2015), 'Ground- and space-based gamma-ray astronomy', *Annu. Rev. Nucl. Part. Sci.* **65**, 245ff.

Gast, H. (2012), A new lookup scheme for hap. H.E.S.S. internal note.

HESS Collaboration, L. (2016), 'Gamma-ray blazar spectra with h.e.s.s. ii mono analysis: The case of pks 2155–304 and pg 1553+113', *A&A* .

Hofmann, W. (2005), 'Hess provides new view of gamma-ray sky', `cerncourier.com/cws/article/cern/29249`. Accessed: 11.03.2017.

J. Hedderich, L. S. (2016), *Angewandte Statistik*, Springer Spektrum.

Lucy, L. B. (1974), 'An iterative technique for the rectification of observed distributions', *The astronomical journal* **79**, 745ff.

O'Haver, T. (2008), 'Fourier deconvolution', `terpconnect.umd.edu/~toh/spectrum/Deconvolution.html`. Accessed: 20.02.2017.

Oxford Dictionary (2017), 'Deconvolution', `https://en.oxforddictionaries.com/definition/deconvolution`. Accessed: 20.02.2017.

Richardson, W. H. (1972), 'Bayesian-based iterative method of image restoration', *Journal of the optical society of America* **62**, 55ff.

University of Nova Gorica (2017), `www.ung.si/en/research/laboratory-for-astroparticle-physics/projects/cta/`. Accessed: 19.01.2017.

Danksagung

An dieser Stelle möchte ich mich bei all denen bedanken, ohne die diese Arbeit nicht zustande gekommen wäre. Insbesondere bedanken möchte ich mich bei

- Prof. Dr. Christopher van Eldik für die überaus interessante Themenstellung, die sehr gute Betreuung und die stets produktiven Diskussionen zu meiner Arbeit
- meinen Zimmerkollegen Andreas, Giacomo und Domenico für die entspannte, aber arbeitsfördernde Atmosphäre, die angeregten Diskussionen und für große Unterstützung beim Erstellen meiner Arbeit
- Johannes für die vielen Erklärungen der Analyse-Software, die stetige Erreichbarkeit bei Fragen sowie die Auswahl und Analyse der H.E.S.S. Phase 2 Daten
- allen Korrekturlesern für ihre Zeit und Nerven, ihre hilfreichen Kommentare und Anregungen und für die positiven Rückmeldungen zu meiner Arbeit
- sowie dem gesamten ECAP für das freundliche Aufnehmen

Erklärung

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

Erlangen, den 13.04.2017

Maximilian Oberndörfer