PROCEEDINGS OF SCIENCE

Towards a 3D analysis in Cherenkov γ-ray astronomy

Jouvin L.[1], **Deil C.**[2], **Donath A.**[2], **Kerszberg D.**[3], **Khelifi B.**[1], **Lemière A.**[1], **Terrier R.**[1], **.**
E-mail: lea.jouvin@apc.in2p3.fr

[1] *APC (UMR 7164, CNRS, Université Paris VII, CEA, Observatoire de Paris), Paris*
[2] *Max-Planck-Institut für Kernphysik, PO Box 103980, D 69029 Heidelberg, Germany*
[3] *LPNHE, Universiteé Pierre et Marie Curie Paris 6, CNRS/IN2P3, Paris*

The development of a 3D or cube analysis to simultaneously fit a spectral and morphological model to the data is a challenge in Cherenkov γ-ray astronomy. The strong variation of the instrument response functions and of the residual cosmic rays background with the observation conditions makes it very difficult to build a coherent background model across the whole energy range. Nevertheless, with improving sensitivities and angular resolution of current Cherenkov telescopes and next generation instruments such as the Cherenkov Telescope Array, the complex morphology of the regions with diffuse emission or multiple sources requires the development of this technique.

We developed a prototype for cube analysis along with corresponding background models for varying observations conditions. The prototype as well as the production of background model was implemented in the software Gammapy, an open source Python package for γ-ray Astronomy that provides tools to simulate and analyse the γ-ray sky for imaging atmospheric Cherenkov telescopes. To validate the method, we performed a systematic comparison between this cube analysis and the classical 1D spectral fitting on point sources using a Monte Carlo simulation from the H.E.S.S. collaboration.

35th International Cosmic Ray Conference — ICRC2017
10–20 July, 2017
Bexco, Busan, Korea

PoS(ICRC2017)765

1. Introduction

With the improvements of the sensibility and angular resolution in Cherenkov very high energy (VHE) γ-ray astronomy, the sky images has revealed sources and diffuse emission with a morphology always more precise and complex. The development of a 3D analysis (also called cube analysis) to simultaneously fit a spectral and morphological model on a data cube (energy, longitude and latitude) is therefore crucial in order to be able to distinguish the several components of a same emission. The 3D analysis requires the knowledge a-priori of a background model. In Cherenkov γ-ray astronomy, instrument response functions (IRF) and the residual background present strong variations with the observation conditions. Is makes it very difficult to build a coherent background model whatever the energy band.

In the section 2, we will present the general principle of the 3D analysis we have implemented in the software Gammapy [1], an open-source Python package for γ-ray astronomy that provides tools to simulate and analyse the γ-ray sky for imaging atmospheric Cherenkov telescopes. In particular, we will briefly explain the method we adopt to construct a background model from H.E.S.S. data. In the section 3, we will present the Monte Carlo (MC) tool we developed to test the 3D analysis. From MC simulation realized by the H.E.S.S. collaboration, a systematic comparison between the 3D analysis and the classical 1D spectral fitting on point sources will be presented.

2. 3D analysis

Principle The 3D analysis is the simultaneous fit of a spectral and morphological model on a data cube. It is based on the comparison between a cube of measured events (ON cube) and a cube of predicted events by a combined model (figure 1). The 3D cube is defined for a given dataset using a 2D spatial binning (e.g. longitude, latitude) and a 1D energy binning. The ON cube is composed of the residual background as well as the γ-ray signal from sources. Therefore, in order to be compared to the data, the predicted data cube have also to be composed of a background (predicted by a background model) and from a γ-ray signal (predicted by the source model).

The adjustment of the spectral and morphological parameters assumed for the source is obtained by a comparison between the number of predicted events by the model and the number of measured events in each energy bin and pixel of the cube. This comparison is realized by minimizing a likelihood function taking into account the Poisson distribution of the number of events in each bin with the Python package Sherpa (CIAO v4.5 [2]) that was adapted in Gammapy for a 3D fitting combining a spatial and a spectral component.

Background Model For the background it is not possible tu use MC simulation, essentially for storage capacity and computing time reasons but also due to the limitation of the hadronic shower models and of the atmosphere knowledge. Our approach to construct the model background cube in reconstructed energy for each run is based on radial acceptance curves. The acceptance tables give the radial dependence of the residual background events. They are determined from region without γ-ray source in the H.E.S.S. observations and assuming a radial symmetry to the center of the camera. We computed these curves for four zenith angle bands ($0-34^\circ$, $34-49^\circ$, $49-61^\circ$ et $61-72^\circ$) and 25 reconstructed energy bins logarithmically spacing between 0.1 and 100 TeV. For high energy bins, high zenith angle bands or large offset, the tables can present strong

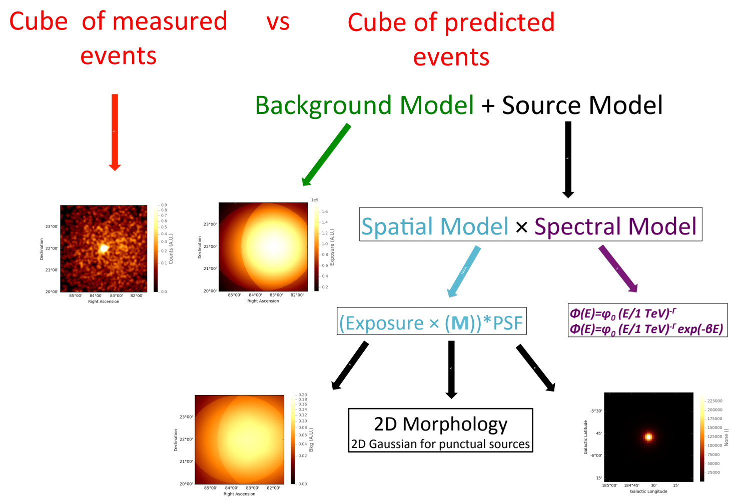


Figure 1: Illustration of the construction of the cube of predicted events for one energy band from the background cube, the exposure and PSF cube. The spectral and spatial parameters are then fitted by comparing the predicted cube events to the measured cube events for each pixel of the images and each energy bin of the cube.

fluctuations. We developed a smoothing algorithm of these curves based on an adaptive gaussian kernel depending on the available statistic used to create the curves.

For a given zenith angle band, the radial acceptance is estimated from the previous model curves by integrating them on the analysis energy bin. For an energy band 1.5-4 TeV, the radial acceptance is presented on the figure 2.a. Assuming a radial symmetry, for a given observation at a certain zenith angle, a model background map is created by rotating the radial acceptance curve for each energy bin of the cube. By summing the background maps of each observation, we obtain a predicted background map for a set of observations. Repeating the operation on each energy band of the cube, we obtain a predicted background cube. The normalisation of the background is done for each energy bin. By using the regions without source in the FOV of the cube (spatial dimension), we impose the number of measured events in these regions to be equal to the number of events predicted by the background model.

Source model The source model is composed of a:

- *spectral model*: spectral law, ϕ, e.g. a power law or a power law with an exponential cutoff.
- *spatial model*: represents the morphology M of the source. For example, for a source with a simple symmetric morphology as we will study in the section 3, we can use a 2D symmetric Gaussian. The adjusted width will constrain the extension of the source, the position of the maximum will give the position of the source and the amplitude will allow to obtain the flux of the source.

Instrument response function The IRFs are computed via MC simulation of γ-ray showers and of the detector. For the 3D analysis, three IRFs are necessary:

- the mean exposure cube (EXPOSURE). In particular for the study of diffuse emission, it is really important to take into account the gradients of the exposure. This is why the spatial model is multiply by the exposure cube. The exposure is the product of the effective area by the time of each observation. The total exposure cube is the sum of the exposure cube for each observation of the given dataset. For a given dataset composed of N_{obs} observations, it is evaluated at each true energy E_{true} at the center of the energy bin of the cube and at each pixel with an offset θ^i_{coord} from the center of the FOV of the run i:

$$Exposure(\theta^i_{coord}, E_{true}) = \sum_{i=1}^{N_{obs}} A_i(\theta^i_{coord}, E_{true}) \times T^i_{obs} \tag{2.1}$$

  with A_i the effective area and T^i_{obs} the livetime associated to the run i.

- the mean point spread function cube (PSF) where the PSF is the spatial distribution of the reconstructed γ-ray directions from a point source. It is a probability. For the H.E.S.S. telescope, a radial symmetry has always be assumed and we only compute 2D table.

- the energy resolution, R, gives the probability to reconstruct an event at a reconstructed energie E' knowing that its true energy is E.

Currently in the 3D analysis in Gammapy, the energy resolution and the PSF are not calculated for each pixel in the FOV of the cube. Both are determined before the analysis at the center of the region of interest (ROI). This approximation implies that they are constant over the ROI. We will have to test the impact of this approximation but it allows a major gain of computation time.

The PSF is determined in function of the distance to the source rad. The mean PSF of the N_{obs} observations is evaluated at each true energy E_{true} of the cube and at a position θ^i_{pos} of the ROI from the center of the FOV of the run i:

$$PSF(rad, \theta^i_{pos}, E_{true}) = \frac{\sum_i^{N_{obs}} A_i(\theta^i_{pos}, E_{true}) \times T^i_{obs} \times PSF_i(rad, \theta^i_{pos}, E_{true})}{\sum_i^{N_{obs}} A_i(\theta^i_{pos}, E_{true}) \times T^i_{obs}} \tag{2.2}$$

It is weighted by the effective area in order to conserve the number of predicted counts when we stack the observations. By assuming a radial symmetry, we project this 1D table on the 2D map for each energy of the cube.

The data cube and the background cube are constructed in reconstructed energy. Consequently, the cube of the predicted events by the source model have also to be converted in reconstructed energy. This is why we have to consider the energy resolution of the instrument. The number of predicted counts N_{true} in a true energy bin E_{true} is given by:

$$N_{true} = (\mathrm{2D\,spatial\,model} \times \mathrm{Exposure(E_{true})}) * \mathrm{PSF(E_{true})} * \phi(\mathrm{E_{true}}) \tag{2.3}$$

By combining the number of predicted counts in each true energy bin with the energy resolution R, we obtain the number of predicted counts in each reconstructed energy bin [${E_{reco}}^i$,${E_{reco}}^{i+1}$] by the following relation:

$$N_{pred}({E_{reco}}^i, {E_{reco}}^{i+1}) = \int_{E_{true}} N_{true} \int_{{E_{reco}}^i}^{{E_{reco}}^{i+1}} R(E_{true}, E_{reco}) dE_{reco}\, dE_{true} \tag{2.4}$$

The mean energy resolution for a given dataset composed of N_{obs} observations is determined similarly to the mean PSF (equation 2.2).

3. First validation of the 3D analysis implemented in Gammapy with a MC tool

In order to validate the 3D analysis approach in Gammapy, we used a MC simulation of γ-ray showers [4] and of the H.E.S.S. detector [5] provided by the H.E.S.S. collaboration. The reconstructed events of these simulations are obtained by applying the reconstruction and γ/background discrimination algorithms from one analysis pipeline of H.E.S.S. [6]. We simulated a test source from the reconstructed events list of these simulations and add these events to the ones of a real H.E.S.S. observation creating thus a fake observation. For each of these observations, we applied the 1D classical spectral analysis and the 3D analysis presented in the previous section, in order to perform a systematic comparison between the two approaches for the reconstruction of the spectral parameters. The 1D spectral fitting implemented in Gammapy is based on the forward folding method and the profile likelihood approach [7, 8] and is similar to the one in the standard H.E.S.S. analysis tools.

We used one MC simulation matching with a zenith angle of 0°, an instrument optical transmission of 80% of the nominal one and an offset of the γ-ray signal to the center of the FOV of 0.5°. The γ-rays are generated following a power law $E^{-\Gamma}$ of spectral index $\Gamma = 2$. In order to simulate a test source in the FOV of one observation, we select randomly and uniformly different samples from the reconstructed events of the MC simulation. These samples match with different statistic representing different signal to noise ratio. In order to test other spectral laws with a spectral index $\Gamma > 2$ or with an energy cutoff, the events are selected from the MC simulation depending on their energy to follow the new spectral law. The number of simulated γ-rays, N_{simu}, is around 760 000 on a total simulated area A with a radius of $R = 550$ m ($\sim 3500\ m^2$). After applying the reconstruction and discrimination algorithms of the used H.E.S.S. analysis methods, the number of reconstructed events, N_{reco}, is around 20 000. For a given observation time, T_{obs}, the different numbers of selected MC events, N_t, match with different flux of the test source, ϕ, estimated by:

$$\phi = (\frac{N_{simu}}{N_{reco}} \times N_t)/(A \times T_{obs})$$

In order to test the sensibility of the 3D analysis to our background model, we use real background events of a run from the observation of the AGN PKS 2155-304. The test source is injected in this run having same observational conditions than the used MC simulation as shown in figure 2.b. We removed all the events from PKS 2155-304 (for safety a region of 0.3° radius). The background events of the real observation are visible throughout the FOV.

We then applied on this fake observation the 1D spectral fitting and the 3D analysis. The general characteristics of the two approaches are summarized in the table 1. For the 3D analysis the total size of the FOV is $2.5° \times 2.5°$ to normalise correctly the background model. The fitting is then realised on an ROI of (1°,1°) around the MC source. For point source, in order to increase the signal to noise ratio, the ROI during the 1D classical spectral fitting is set to 0.1°. Specific IRFs are computed for this analysis taking into account the angular cut applied on the events. For the 3D analysis, the source is assumed to be fully contained in the ROI.

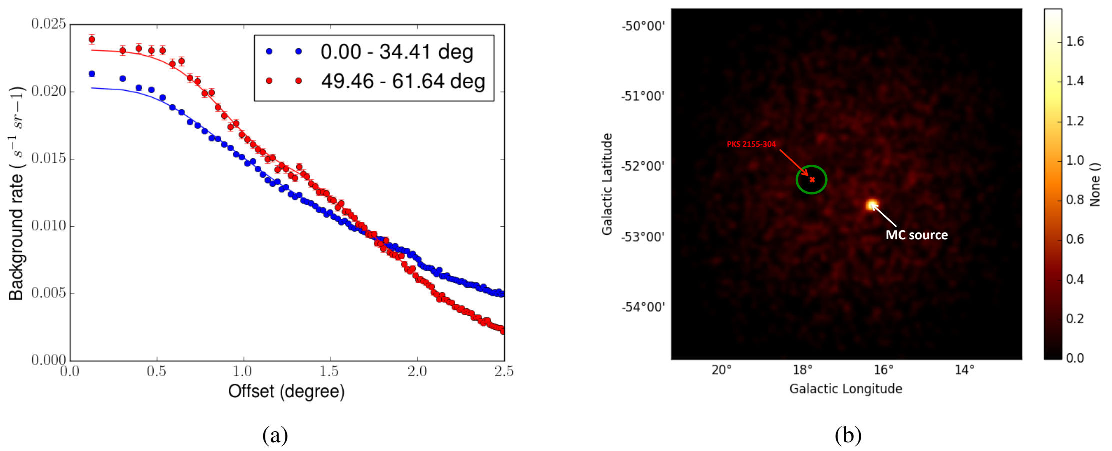


Figure 2: (a) Background rate depending on the offset to the center of the camera for an energy band 1.5-4 TeV and two zenith angle bands: 0-34° (blue) and 49-61° (red). The solid lines are the smooth acceptance curve (b) Spatial distribution of the simulated events for the MC source (constituted of 100 events) and of the background events from an observation of the AGN PKS 2155-304.

Table 1: Characteristics of the 1D and 3D analysis

1D Analysis	3D Analysis
- E=0.1-100 TeV, 20 bins - fitting energy range: determined on the available statistic in each energy bin - ROI: disk centered on the MC source with R=0.1° - Background computation: reflected region [3]	- Cube: centered on the MC source - E=0.5-100 TeV, 20 bins - FOV: 2.5° $\times$ 2.5° - spatial morphology: 2D symmetric Gaussian - ROI=(1°,1°) around the MC source - Background computation: from our background model

For each value of N_t of MC γ-ray injected, we produced 200 sets of randomly selected N_t events. For each one, we applied the two fitting procedures. Finally we obtained a distribution of the reconstructed parameters from these 200 test observations. On the figure 3, the blue points represent the median of the distribution and the error bars the dispersion around the median (interval around the median containing 68% of the distribution). For the spectral component, we have simulated two spectral laws: a simple power law with a spectral index $\Gamma = 2$ (figures 3 a,b) and a power law with an exponential cutoff at 10 TeV for a spectral index $\Gamma = 2$ (figures 3 c,d). The spectral parameters on the panels on the left are adjusted with the 1D spectral fitting and on the panels on the right with the 3D analysis.

For the two spectral laws, the spectral parameters and in particular the exponential cutoff are well reconstructed when the statistic is sufficient. The median of the distribution is well centered on the injected parameters. The bias and the error bars are larger for the 1D spectral analysis since for a given sample size, the ROI used for the 1D spectral fitting is 0.1°. Only $\sim$ 50% of the total events sample are used in the fit. For the low statistic, several observations are indeed necessary to detect significantly a source. For an observation of around 28 minutes, the 1D spectral fitting presents a bias that was already observed in the standard H.E.S.S. analysis tool. In the 3D analysis, the fit is realized on a (1°,1°) region around the MC source. Therefore all the simulated events are used during the analysis. This is why, even with only 100 events, we observe very small bias with the 3D analysis. We also tested softer spectral laws and higher energy cutoff. The spectral parameters were still perfectly reconstructed, in particular with the 3D analysis. The bias in the 3D analysis, inferior to 1% , is small. It is well below the CTA requirements for other systematics on parameters produced by early analysis steps, e.g. calibration, statistical error on the effective area. If the 3D analysis bias stay as small as presented here, the main systematics for CTA will come for our limited knowledge of the instrument.

In the 3D analysis, in addition to the spectral components, we simultaneously fitted a morphology for the source. In this test, we used a 2D symmetric Gaussian. The width of the Gaussian allows to estimate the constraints we have on the source extension, being here inferior to 14 arcsec taking 90% of the distribution values. The simulated MC source is then well compatible with a point source. In parallel, the position of the Gaussian is always adjusted on the real source position. The error we get on the source position of 0.01° for a spatial binning of the cube of 0.01° comes from the intrinsic method we are using to convolve by the PSF cube. Indeed the PSF cube has the same spatial binning than the data cube. Therefore the position of the center of the 2D gaussian is limited by the spatial grid. For point source, a more precise method must be developed for the convolution by the PSF. One idea could be to use a thinner spatial binning for the model cube. The 3D analysis is perfectly able to reconstruct the spectral and spatial parameters of the point source.

4. Conclusion and Perspectives

This new MC tool allows to inject a fake MC source in an observation. This is the first step for validating the 3D analysis recently implemented in Gammapy. For one observation, we validated the reconstruction of the spectral parameters for a power law or a power law with an exponential cutoff for different flux of the MC source. The reconstruction of the spectral and spatial parameters, simultaneously adjusted, is well under control with small bias.

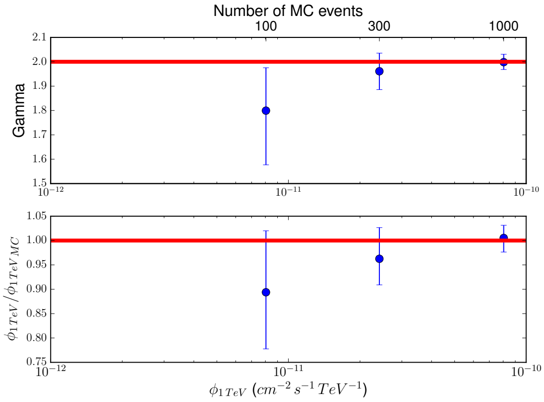
The future tests will be first to simulate a MC source in an observation with different observational conditions. To test the adjustment of the morphology, we have to create morphologies for more extended sources. Finally, to test the reconstruction of the spectral and spatial parameters for lower statistics, we have to develop the simulated data in multiple observations with different observational conditions stacked together. We will then be able to check the stacking of the observations and the computation of the mean instrument functions.

Acknowledgement

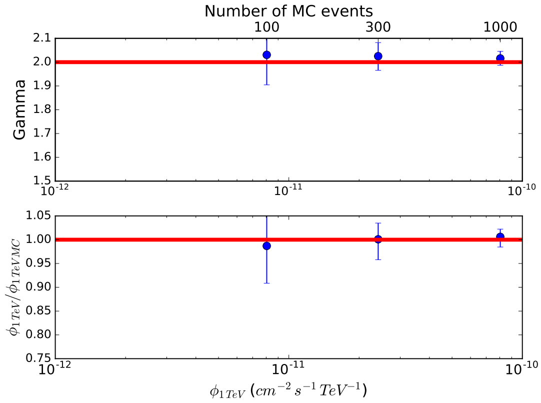
We are really grateful to the H.E.S.S. collaboration for providing us the H.E.S.S. observations to build the acceptance curves and the MC simulations of the γ-ray showers and of the detector.

References

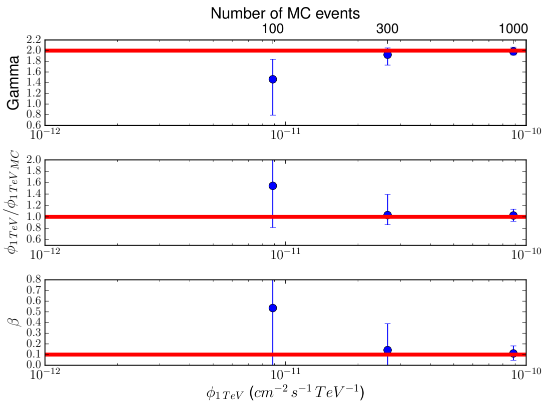
[1] Deil C. et al, 2017, Proceedings of the 35th International Cosmic Ray Conference



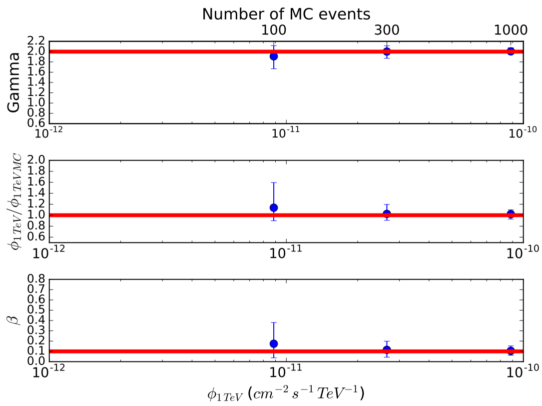
(a) 1D Analysis, $\Gamma = 2.0$



(b) Analyse 3D Analysis, $\Gamma = 2.0$



(c) 1D Analysis, Γ=2.0



(d) 3D Analysis, Γ=2.0

Figure 3: Median of the spectral parameters distribution fitted on 200 test observations (blue points). The MC source is simulated with a simple power law ($\propto E^{-\Gamma}$) of spectral index $\Gamma = 2$ (a,b) and with a power law with an exponential cutoff ($\propto E^{-\Gamma} exp(-\beta E)$) at 10 TeV ($\beta = 0.1$) and of spectral index $\Gamma = 2$ (c,d) for different differential flux at 1 TeV (abscissa on the figures). The errors bars match the dispersion of the distribution around the median value. The red horizontal lines are the injected parameters used to simulate the MC source. On the left figures, the results are obtained using the 1D classical spectral analysis and on the right plots using the 3D analysis.

[2] Refsdal B. et al, 2009, Proceedings of the 8th Python in Science Conference

[3] Berge D. et al, 2007, A&A, 466, 1219

[4] Kertzman M. P. & Sembroski G. H., 1994, Nuclear Instruments and Methods in Physics Research A, 343, 629

[5] Guy J., 2003, PhD Thesis

[6] Khelifi B. et al, 2015, Proceedings of the 34th International Cosmic Ray Conference, 837

[7] Piron F. et al, 2001, A&A, 374, 895-906

[8] Rolke W., Lopez A.M. and Conrad J., 2005, Nuclear Instruments and Methods in Physics Research A, 551, 493-503