

Statistical biases of spectral analysis with the ON-OFF likelihood statistic

Jouvin L.*¹, Lemière A.¹, Terrier R.¹, Ohm S.², Oya I.², van Eldik C.³.

E-mail: lea.jouvin@apc.in2p3.fr

¹ APC (UMR 7164, CNRS, Université Paris VII, CEA, Observatoire de Paris), Paris

² DESY, D-15738 Zeuthen, Germany

³ Universität Erlangen-Nuremberg, Physikalisches Institut, Erwin-Rommel-Str. 1, D 91058 Erlangen, Germany

For the H.E.S.S. Collaboration

Spectral extraction in data from ground-based gamma-ray observations is usually performed using the on-off likelihood statistic which is based on the profile likelihood technique. The latter is known to lead to inconsistent estimators in some situations. We present here a systematic Monte Carlo study of the distribution of fitted spectral parameters for typical observations with VHE (very high energy; >100GeV) observatories and show that in some conditions spectral extraction yields inconsistent estimators. We discuss some techniques to alleviate this effect and the impact on the search for spectral cutoffs in the very low statistic regime.

*The 34th International Cosmic Ray Conference,
30 July- 6 August, 2015
The Hague, The Netherlands*

*Speaker.

Introduction

Spectral analysis of cosmic γ -ray sources is complex in very high energy (VHE; 100 GeV-100 TeV) astronomy. Indeed, because of the low signal to noise ratio the reconstruction of the spectral parameters of a power law for faint sources is not simple. The reconstructed spectrum can be softer than expected. The detection of a spectral feature such as a cutoff has strong physical implications. This is why we have to test the reliability of the fitting procedure applied on the VHE data based on a "forward-folding" method [1].

This method is based on a comparison of a spectral shape hypothesis assumed a priori convolved by the instrument response functions (IRF) with the observations. The parameters of the a-priori chosen shape hypothesis for the source are then evaluated by a maximum likelihood (ML) method taking into account the energy distribution of the number of γ -ray candidates in the ON and OFF regions that follow Poisson distributions [2].

With the H.E.S.S. telescope array, observing the sky in the VHE domain, the dataset of a given region consists of various observations that match with a set of observational parameters. For each observation of a field of view, we define a region of interest including the studied source (ON region) and a region devoid of sources (OFF region) allowing to estimate events caused by the residual hadronic background in the ON region. In order to avoid the use of a model for the background, hard to construct because of varying observing conditions, we resort to a modified ML method named profile likelihood [3]. The "profile likelihood" method introduces the background rate estimators for each reconstructed energy bin. A likelihood function taking into account the poissonian probability to obtain the measured ON and OFF events for all the observations of the studied source or region is built in order to determine the parameters of the supposed spectral shape. By maximizing the likelihood function over the background, the estimator of the background rate can be written as a function depending only on the signal rate estimator [4].

As explained in [5], when the number of nuisance parameters is increasing with respect to the number of observations, the profile maximum likelihood estimators are likely to be inconsistent (i.e. do not converge to the right value even for an infinite number of observations). Without any models for the background, we have to measure an OFF region for each observation. Therefore, there can be as many nuisance parameters as observations, and the number of degrees of freedom can increase with the number of observations and energy bins used for the maximisation of the likelihood function in the profile likelihood method.

In order to keep all the information related to the observation conditions, one could design a spectral analysis that processes all the observations individually for the maximisation of the likelihood function. However, except for the observations with a high signal to noise ratio, this procedure leads to inconsistent estimators when the number of observations becomes large. In order to limit the number of degrees of freedom in the fit and to increase the statistics, it is thus necessary to group observations together in bands of similar observational conditions. We developed a Monte Carlo (MC) tool that allows to perform systematic studies on simulated spectra, using one of the official H.E.S.S. spectrum fitting packages, to test and evaluate the biases of the estimators obtained with this method on the adjusted spectral parameters. We test the impact of grouping the observation together on the estimators bias.

1. Spectrum simulation

An IACT (Imaging Array Cherenkov Telescope) study of a particular γ -ray source is typically composed of several observations (corresponding to a set of observational parameters: zenith angle Z , offset in the field of view θ and optical efficiency ε). Each observation is split in several energy bins. The IRFs have a strong dependence with observational parameters and energy and are therefore calculated for each observation. For each observation j and each energy bin i , we measure the number of ON and OFF events ($N_{ONi,j}$ and $N_{OFFi,j}$) as well as the exposure normalization between the OFF and ON region α_j .

We developed a MC tool within one of the H.E.S.S. spectrum fitting packages "START" [7] that provides a simulated number of "measured" ON events and OFF events in each energy bin for a given observation and source spectral shape taking into account the response of the H.E.S.S. instrument to gamma-ray events. For a set of observation, in order to test the validity of the adjusted spectral parameters, we can apply the fitting procedure to these simulated data.

For each observation j and energy bin i , the expected number of events in the ON data $\overline{N}_{i,j}$ is composed of the signal coming from the source $\overline{S}_{i,j}$ and the number of events expected in the OFF data $\overline{B}_{i,j}$:

$$\overline{N}_{i,j} = \overline{S}_{i,j} + \alpha_j \overline{B}_{i,j}$$

where α_j is the normalization between the ON and the OFF regions for each observation j .

Signal

In this study, we model the differential flux of the simulated source by simple laws based on power laws. The expected number of events coming from the source results from a convolution between the intrinsic source spectrum and the IRFs. The number of events measured in a reconstructed energy bin i [$E_{m_i}, E_{m_{i+1}}$] for an observation with a lifetime ΔT is:

$$\overline{S}_{i,j} = \Delta T \int_{E_{m_i}}^{E_{m_{i+1}}} \int_{E=0}^{\infty} \phi(E) \times IRF(E, E_m, Z, \theta, \varepsilon) dE dE_m$$

where $\phi(E)$ is the expected differential flux of the source.

For each observation and energy bin, the signal events in the ON data are generated following a Poisson distribution $P(\mu_1)$ with the expectation value $\mu_1 = \overline{S}_{i,j}$.

Background

The OFF data for each observation j and energy bin i represent a certain realisation of the expected number of background events. However, for energies exceeding 1 TeV, the number of background events in any energy bin is typically low. Therefore for these energies we model the background by a power law of spectral index 2.7¹. We require that the total number of expected events above 1 TeV from the model matches that of the recorded background events. For each observation j and energy bin i , we have:

¹We tested also spectral indices of 2.6 and 2.5 and different energy thresholds and found no impact on the reconstructed parameters

$$\begin{cases} \overline{B}_{i,jE < 1 \text{ TeV}} = N_{OFF i,j} \\ \overline{B}_{i,jE > 1 \text{ TeV}} = \int_{E_i}^{E_{i+1}} \phi_j \times \left(\frac{E}{1 \text{ TeV}} \right)^{-2.7} dE \end{cases}$$

with: $\int_{1 \text{ TeV}}^{E_{max}} \phi_j \times \left(\frac{E}{1 \text{ TeV}} \right)^{-2.7} dE = \sum_{1 \text{ TeV}}^{E_{max}} N_{OFF i,j}$

We use the OFF data to simulate the background in the ON data. The number of background events in the ON data are distributed according to a Poisson distribution $P(\alpha_j \overline{B}_{i,j})$. The number of events in the OFF data are also generated according to a Poisson distribution of parameters $\overline{B}_{i,j}$.

To test the consistency of the spectrum model parameters, we simulated 150 γ -ray spectrum, where we used 443 observations taken by H.E.S.S. between 2004 and 2012 matching with the observations of the Galactic Center [6]. The data set comprises 250 hours of observations of the GC observed many times since 2004, taken at a mean zenith angle of 22° ($Z \in [6^\circ; 60^\circ]$), offset of 0.8° ($\theta \in [0.2^\circ; 2.2^\circ]$) and optical efficiency of 60% ($\epsilon \in [50\%; 80\%]$). Because of the large variation in observation parameters, this data set is an ideal test case for the systematic studies that we want to perform.

2. Results for the power law

2.1 Evolution of the inconsistency with flux

The figure 1 represents the distribution of the fitted differential flux normalization parameters ϕ_0 at 1 TeV and the spectral index Γ . We define the unit ϕ_{0Crab} matching with the Crab-Nebula differential flux normalization at 1 TeV ($\phi_0 = 4.44 \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1} \text{ TeV}^{-1}$). For a bright source, $\phi_0 = \phi_{0Crab}$, (figure 1 (a) and (b)), the distributions of the adjusted spectral parameters are well centered on the input values used in the simulation. For a very faint source, $\phi_0 = 0.5\% \phi_{0Crab}$, (figure 1 (c) and (d)), the simulation clearly shows that the spectrum estimators are biased, since there is a systematic difference between the mean reconstructed values and the input value (the gap can reach dozens of percent) even if the total significance is around 14σ and the total excess around 550. We interpret this effect as a consequence of the large amount of nuisance parameters (background) in the case of low-statistic.

In the following sections, we study the impact of combining the observations that is often used to improve the statistic at high energy.

2.2 Impact of the observations grouping

For faint sources with a large data set, the low signal to noise ratio can lead to inconsistency in the spectral fitting procedure. In order to remove this issue, it is necessary to group several observations together to improve the statistics per energy bin and limit the nuisance parameters. For different observational parameters, the IRFs and the OFF events distribution can present large variations. Therefore, the observations with similar observational parameters are combined in different bands of zenith angle, offset and efficiency. The mean zenith angle $\langle Z \rangle$, the offset $\langle \theta \rangle$ and

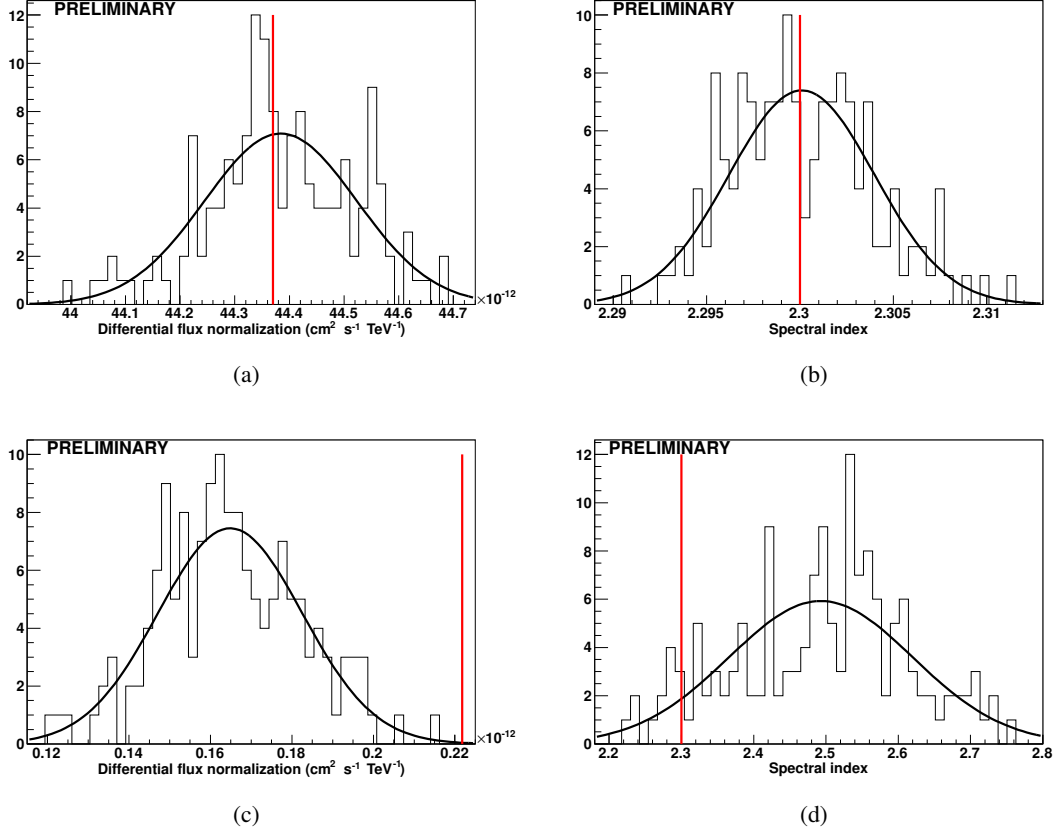


Figure 1: Distribution of the differential flux normalization at 1 TeV ϕ_0 and the spectral index Γ for 150 simulated spectra with a power law ($\phi_0 (E/1 \text{ TeV})^{-\Gamma}$) of spectral index $\Gamma=2.3$ and a differential flux normalization at 1 TeV $\phi_0 = \phi_{0\text{Crab}}$ in (a) and (b) and $\phi_0 = 0.5\% \phi_{0\text{Crab}}$ in (c) and (d). For each observation, the data are simulated in 12 energy bins between 0.25 and 30 TeV. The red vertical lines represents the injected parameters used to simulate the MCs. Each distribution is fitted with a Gaussian model using a maximum likelihood method.

the efficiency $\langle \epsilon \rangle$ of each band are calculated as the mean of the parameters of each observation in the band, weighted by the livetime of each observation. The IRFs of each band are then evaluated for $\langle Z \rangle$, $\langle \theta \rangle$ and $\langle \epsilon \rangle$. The likelihood function is then maximized band by band and energy bin by energy bin. For our set of observations (443), the runs grouping leads to 43 bands.

In the following, we test the impact of the observations grouping on the performance of the estimators. The MC simulation is carried out observation by observation, then the observations are grouped together and the spectrum is fitted band by band.

The figure 2 represents the mean value of the adjusted parameter distributions (blue points) as a function of the true flux (used to simulate the MC). As observed on the figure 2(a), the fitted spectrum becomes softer as the differential flux normalization decreases. For the high differential flux values, the reconstructed value for Γ and ϕ_0 are not biased. However, from a differential flux normalization $\phi_0 = 1\% \phi_{0\text{Crab}}$, without combining the observations, the adjusted value of ϕ_0 and Γ are not in agreement with the injected value even if the significance is around 15σ and the total excess around 1000. The bias is larger than the typical statistical error on the fit and reaches, for the

lowest differential flux, around 20%. For harder power-law the bias decreases in particular on the Γ parameter (around 8% for the lowest differential flux normalization for a power law of photon index $\Gamma=1.6$) and increases for a softer one (around 30% for a power law of photon index $\Gamma=2.6$).

The figure 2(b) presents the results of the previous MC test when the observations are grouped together. Combining the observations in different bands of zenithal angle, offset and efficiency improves the quality of the adjustment of ϕ_0 and Γ for the low differential flux values by improving the statistics at high energy (only a few percent bias for the lowest flux).

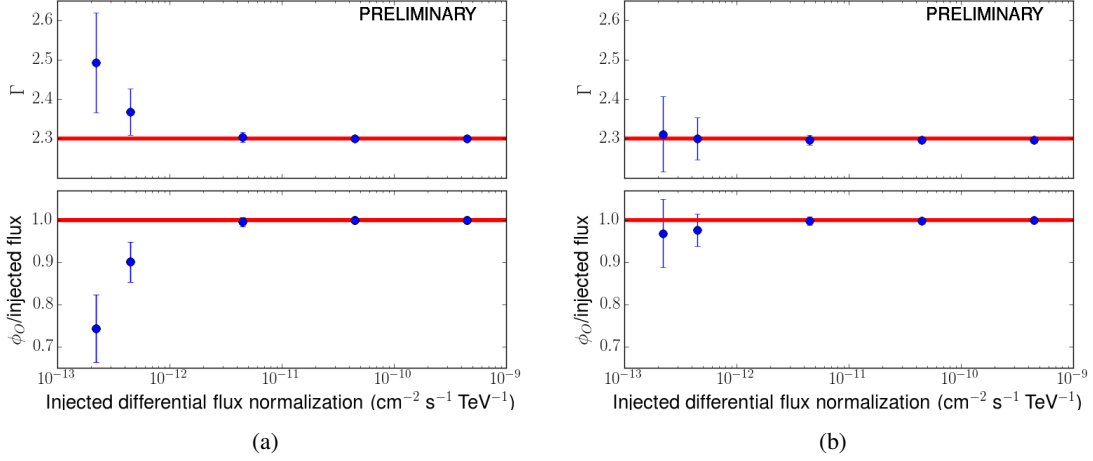


Figure 2: Mean value of the adjusted parameters distribution on 150 simulated spectra (blue points) generated with a power law of spectral index $\Gamma=2.3$ and for different differential flux normalization values at 1 TeV equal to 0.005, 0.01, 0.1, 1 and 10 ϕ_{0Crab} (abscissa on the plots) without any observations grouping in (a) and with the observations grouping in (b). For each observation, the data was simulated in 15 energy bins between 0.25 and 30 TeV. The error bars are the standard deviation of the distributions. The red line is the injected spectral parameter of source.

For low fluxes ($\sim 1\% \phi_{0Crab}$), a residual bias introduced by the "profile likelihood" method is clearly visible for the spectral parameters of a simple power law. In the next section, we test the reconstruction of the cutoff energy by generating the MCs by a power law with an exponential cutoff.

3. Result for the power law with an exponential cutoff

The signal is generated by a power law with an exponential cutoff for different cutoff energies (see details figure 3). The figure 3 represents the mean value of the adjusted parameters as a function of $\beta = \frac{1}{E_{cut}}$: in (a) without grouping the observations and in (b) with the observation grouping.

Due to the low statistics, the spectral parameter β is always higher (so the cutoff energy always lower) than the one injected to simulate the signal in the ON data (around two times lower for the highest cutoff energy of 40 TeV). Due to the correlation between the spectral parameters, the

spectral index is always lower and follows exactly the inverse trend of β . As expected, the cutoff at low energies is better reconstructed than the ones to high energies.

Without any observation grouping, the adjusted value of β and Γ are not in agreement with the injected value. The bias on the three spectral parameters appears at higher flux than for the power law model for which one at an input flux of 5% $\phi_{0_{Crab}}$, the bias is very small even with a per-observation likelihood maximisation. Obviously, grouping the observations together helps to better reconstruct the cutoff energy as well as the Γ and the ϕ_0 (figure 3.b) but the reconstruction of β is still not perfect for the flux of this source (compatible with the injected value taking into account the statistical error but with a residual bias around 10%).

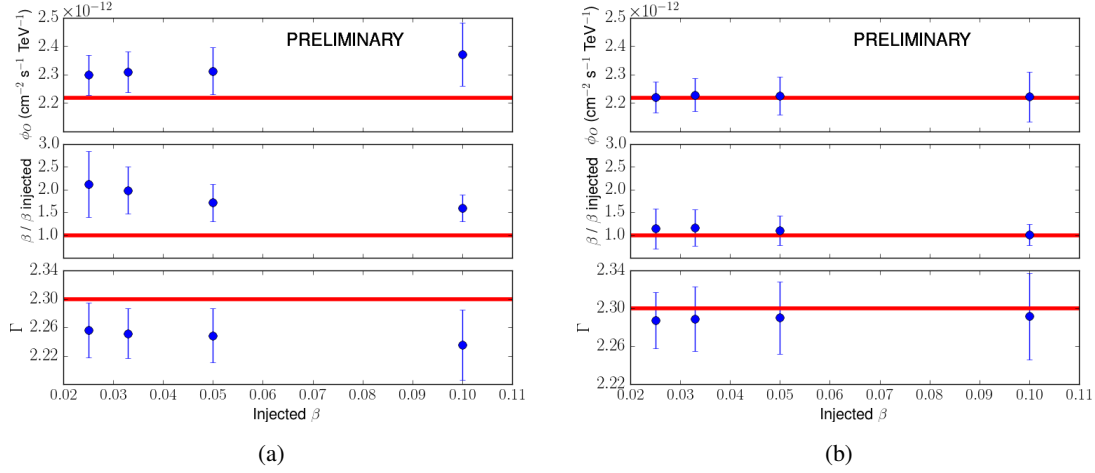


Figure 3: Mean value of the adjusted parameters distribution on 150 simulated spectra (blue points) generated with a power law with an exponential cutoff ($\phi_0 (E/1 \text{ TeV})^{-\Gamma} \exp(E/E_c)$) of spectral index $\Gamma=2.3$, differential flux normalization at 1 TeV=5% $\phi_{0_{Crab}}$ and for different parameters $\beta = \frac{1 \text{ TeV}}{E_{cut}}$ (abscissa on the plots) where E_{cut} is the cutoff energy: $E_{cut}=10, 20, 30,$ and 40 TeV equal to $\beta = 0.10, 0.05, 0.036$ and 0.025 without grouping the observations in (a) and with the observations grouping in (b). For each observation, the data was simulated in 12 energy bins between 0.25 and 50 TeV. The error bars are the standard deviation of the distributions. The significance is around 90σ for the the different cutoff energies and the total excess around 5000. The red line is the injected spectral parameter of the source.

The value of the differential flux for which the spectral parameters are biased depends on the spectral index (for harder spectra the bias decreases) and the number of observations used in the MC simulation.

One hypothesis concerning the increase of the bias for faint sources is the low statistic at high energy. Some ON events are misinterpreted as background events leading to a softer spectrum or to a lower exponential cutoff. Obviously, as for the power law, by increasing the statistic especially to high energy, the observations grouping helps to better reconstruct the cutoff energy as well as the Γ and the ϕ_0 .

4. Conclusions

The MC tool developed allows to simulate ON and OFF observations for a given set of ob-

servations using a source spectral model and the instrument responses. By applying the fitting procedure to the simulated MC spectra, it is then possible to compare the distribution of fitted parameters to the input parameters and look for possible bias. It can also be used to evaluate the sensitivity to a spectral feature (e.g. probability to miss a cutoff or to have a false cutoff detection).

We first showed that likelihood maximization procedure applied observation by observation and energy bin by energy bin leads to very strong biases in the estimators in certain cases, especially for faint sources, for which the background is much stronger than the signal. In this case, the bias on the reconstructed spectral parameters on the MCs can reach several tens of percent.

We then implemented the observations grouping that samples the observations in different bands of zenith angle, offset and efficiency in order to increase the statistics for the faint sources in particular at high energy. We analysed the impact of this grouping on the estimators performances. For power law or power laws with an exponential cutoff fits, combining the observations does improve the reliability of the fitted parameters at low flux values (down to scales of 1% $\phi_{0_{Crab}}$) without introducing large errors. However, even when we combine the observations together, a small bias remains for the reconstruction of the cutoff.

Spectral fits using ON-OFF likelihood can lead to biased and even inconsistent estimators at low source fluxes. The usage of MC simulations is an excellent way to test the consistency of the fitted spectral parameters.

References

- [1] Piron F., A. Djannati-Atai, M. Punch et al, 2001, A&A, 374, 895-906
- [2] Cash W., 1979, ApJ, 228, 939-947
- [3] Neyman J. and Scott E.L., 1948, The Econometrica society, 16, 1-32
- [4] Rolke W., Lopez A.M. and Conrad J., 2005, Nuclear Instruments and Methods in Physics Research A, 551, 493-503
- [5] Spanos A., 2013, ArXiv e-prints, 1301.0118
- [6] Lemière A. et al, Proceedings of the 34th International Cosmic Ray Conference, 2015
- [7] Khelifi B. et al, Proceedings of the 34th International Cosmic Ray Conference, 2015