PROCEEDINGS OF SCIENCE

Bayesian approach to signal estimation in gamma-ray astronomy with Gammapy

Matheus Genaro Dantas Xavier,[a,*] Tim Unbehaun,[a] Rodrigo Guedes Lang[a] and Stefan Funk[a]

[a] *Erlangen Centre for Astroparticle Physics (ECAP), Friedrich-Alexander-Universität Erlangen-Nürnberg Nikolaus-Fiebiger-Str. 2, 91058 Erlangen, Germany*

E-mail: matheus.genaro@fau.de, tim.unbehaun@fau.de, rodrigo.lang@fau.de, s.funk@fau.de

A fundamental challenge for observations with Imaging Atmospheric Cherenkov Telescopes is the treatment of the dominant background of cosmic-ray initiated air showers. Traditional methods for signal estimation rely on gamma-hadron separation cuts to remove a large fraction of background events, thereby reducing the efficiency of gamma-ray detection. In this work we adopt and extend a method known as Bayesian Analysis including Single-event Likelihoods (BASiL), in which these separation cuts are avoided by including into the likelihood function the probabilities of observed events originating from gamma or cosmic rays. With this approach, we retain 10 to 20% more signal throughout the analysis, particularly improving low signal-to-noise regimes and lowering the energy threshold. From the posterior probability of the source signal, credible intervals can be used to derive statistical uncertainties on flux points, while the computation of the Bayes factor allows the assessment of the probability of source detection. By adapting the open-source package Gammapy, we combine BASiL and the forward folding method to fit models and extract source spectra even from highly background-dominated data. This has potential applications to weak-source analyses or transient phenomena. We present results from simulated sources in the context of the H.E.S.S. telescopes, both in the 1-dimensional (data binned in energy) and a novel 3-dimensional analysis extension (data binned in energy and spatial coordinates), investigating how the method performs under distinct observation conditions and source characteristics.

39th International Cosmic Ray Conference (ICRC2025)
15–24 July 2025
Geneva, Switzerland

ICRC 2025
The Astroparticle Physics Conference
Geneva July 15-24, 2025

*Speaker

PoS(ICRC2025)655

1. Introduction

The observation of gamma-ray sources with Imaging Atmospheric Cherenkov Telescopes (IACTs) has the inevitable challenge of dealing with an overwhelmingly dominant background of cosmic rays. Both highly energetic photons and charged particles initiate a cascade of interactions, including particle production in the atmosphere, and develop an extended air shower, in which relativistic charged particles can produce Cherenkov radiation when travelling faster than the speed of light in the medium. Due to the different nature of interactions depending on the primary particle, the development of the shower and consequent emission of Cherenkov light are affected, resulting in distinct image features between gamma rays and hadrons, as recorded by the IACT camera, which allow them to be classified.

Currently, these image features, parametrized by the Hillas variables or other quantities, are provided to machine learning algorithms that perform the gamma-hadron separation. In the current generation of IACTs, we see examples such as the use of random forests for the MAGIC telescopes [1], or boosted decision trees (BDT) [2] in case of H.E.S.S. [3, 4]. From the classification score of these methods, a cut is applied to the reconstructed event data, removing a large fraction of the background. Depending on how strict the cut is, more or less signal events are also discarded, reducing the efficiency of gamma-ray detection. However, even after cuts, an expressive amount of hadronic gamma-like events is still present, requiring background estimation techniques to further identify the gamma-ray signal in the context of a statistical inference analysis.

The standard statistical analysis is based on maximum likelihood estimation and likelihood-ratios for hypothesis testing, as widely known from the works of Li & Ma [5] and Cash [6]. This frequentist approach also depends on asymptotic properties, which are valid, as a consequence of the Poisson nature of the data, in regimes with a sufficiently high number of counts. The alternative Bayesian thinking presents a general approach that has no restrictions on the number of counts, being particularly useful for weak signals [7]. An extension of the Bayesian approach was presented in [8], which provides promising applications to gamma-ray astronomy. In their work, the Bayesian Analysis including Single-event Likelihoods (BASiL) is proposed to perform the reconstruction of the signal without applying gamma-hadron separation cuts. In place of cutting away events, the information from gamma-hadron separation variables is included in the likelihood, effectively weighting all events according to their signal- or background-likeness. With the added information in the construction of the posterior probability distribution of the signal, the method achieves overall lower uncertainties, in particular at low signal-to-noise ratio.

In this work, we implement BASiL in the framework of Gammapy [9], proposing a practical study in the context of the H.E.S.S. telescopes, using simulated data. We show the improvements in statistical uncertainty when reconstructing spectral parameters and also present a new extension of BASiL to a 3D-type analysis, allowing a proper treatment of spatial degrees of freedom. Section 2 describes the methodology and practical implementation in Gammapy, while section 3 presents a comparison between the standard and BASiL analyses in the On/Off approach, as well as a validation of the 3D case. Section 4 discusses some conclusions and perspectives.

2. Methodology

The statistical analysis of reduced gamma-ray data – measured counts binned in spatial and spectral coordinates – depends on the assumptions made regarding the background. When the background is totally unknown (no model parametrization available) and is being directly estimated from data, a class of On/Off estimation methods can be employed, such as reflected regions or ring background [10]. In this case, an On region is defined around the target source, while Off regions in the same field of view (FoV) are used to provide the background estimation, leading to a likelihood, for each energy bin, corresponding to the product of two Poisson distributions

$$L(N_{\mathrm{on}}, N_{\mathrm{off}}|\mu_{\mathrm{s}}, \mu_{\mathrm{b}}, \alpha) = \mathrm{Pois}(\mu_{\mathrm{s}} + \alpha\mu_{\mathrm{b}}|N_{\mathrm{on}}) \times \mathrm{Pois}(\alpha\mu_{\mathrm{b}}|N_{\mathrm{off}}), \tag{1}$$

where N_{on} and N_{off} are the measured counts in the On and Off regions, μ_{s} and μ_{b} the expected signal and background counts, while α is the normalization between both regions that takes into account their different exposure and acceptances [10]. This is the so-called "1D analysis" (counts spatially integrated in these regions).

In the frequentist – hereinafter "standard" – approach, the background can be estimated from the profile likelihood: choosing μ_{b} such that L is maximized, given the signal μ_{s}. As presented by Li & Ma (1983), statistical testing can be performed from the likelihood ratio of competing hypothesis. Given Wilk's theorem [11], valid at the asymptotic limit of high counts, the logarithm of the likelihood ratio is connected to a chi-squared distribution, meaning that significance levels can be computed from the test statistic.

In the Bayesian approach, the goal is to describe the probability distribution of the signal, $p(\mu_{\mathrm{s}}|N_{\mathrm{on}}, N_{\mathrm{off}}, \alpha)$. From Bayes' theorem, the posterior probability is proportional to the likelihood times the prior distribution, while μ_{b} can be "integrated-in" as

$$p(\mu_{\mathrm{s}}|N_{\mathrm{on}}, N_{\mathrm{off}}, \alpha) \propto \int \mathrm{d}\mu_{\mathrm{b}} L(N_{\mathrm{on}}, N_{\mathrm{off}}|\mu_{\mathrm{s}}, \mu_{\mathrm{b}}, \alpha) p(\mu_{\mathrm{b}}) p(\mu_{\mathrm{s}}), \tag{2}$$

revealing the connection to the likelihood of eq. 1. For constant priors, this integral can be analytically performed and the results can be found in [7, 8].

The Bayesian analysis including single-event likelihoods (BASiL) method was proposed by [8] as an extension of eq. 2, by inserting the event characteristics as additional information in the posterior probability. This takes the form of a list of variables x_i for each event in the On region, summarized by $X = \{x_i, 1 \leq i \leq N_{\mathrm{on}}\}$. The calculation of $p(\mu_{\mathrm{s}}|X, N_{\mathrm{on}}, N_{\mathrm{off}}, \alpha)$ then introduces an additional term in the likelihood, associated with the probabilities of observing the values x_i, consequently weighting the events according to the likelihood of having gamma or background origin. The result [8], valid for On/Off analyses, is

$$p(s|X, N_{\mathrm{on}}, N_{\mathrm{off}}, \alpha) \propto \sum_{N_{\mathrm{s}}=0}^{N_{\mathrm{on}}} \frac{(N_{\mathrm{on}} + N_{\mathrm{off}} - N_{\mathrm{s}})!}{(N_{\mathrm{on}} - N_{\mathrm{s}})!(1 + 1/\alpha)^{-N_{\mathrm{s}}}} \frac{C(X, N_{\mathrm{s}})}{\binom{N_{\mathrm{on}}}{N_{\mathrm{s}}}} \frac{\mu_{\mathrm{s}}^{N_{\mathrm{s}}}}{N_{\mathrm{s}}!} e^{-\mu_{\mathrm{s}}}, \tag{3}$$

where $C(X, N_{\mathrm{s}})$ corresponds to the probability of having N_{s} signal events in the On region, given the event variable list X. This requires attributing for each event the likelihoods $p(x_i|\gamma)$ and $p(x_i|\bar{\gamma})$ of observing x_i values, when the event is signal, or has background origin, respectively. The practical computation of $C(X, N_{\mathrm{s}})$ is described in [8].

When the background is known or can be modelled by a template, the likelihood per bin is described by a single Poisson distribution $\mathrm{Pois}(\mu_\mathrm{s}+\mu_\mathrm{b}|N)$, given N measured counts. Following the maximum-likelihood procedure and applying the logarithm of the likelihood ratio, the Cash statistic can be obtained, from which the significance is estimated. With this "3D analysis", significance maps can be produced and the source extension or morphology becomes accessible.

In this work, we propose a simple extension of the BASiL methodology to the 3D case. Considering once more Bayes' theorem for μ_s, the posterior probability per bin becomes

$$p(\mu_\mathrm{s}|N, X) = \int \mathrm{d}\mu_\mathrm{b} \mathrm{Pois}(\mu_\mathrm{s} + \mu_\mathrm{b}|N) p(X|N, \mu_\mathrm{s}, \mu_\mathrm{b}) p(\mu_\mathrm{b}) p(\mu_\mathrm{s}), \tag{4}$$

including the event variables X and separating the terms in the likelihood. We adopt the same constant prior for μ_s as [8], but we now consider the background to be fixed by a model $\tilde{\mu}_\mathrm{b}$, implying

$$p(\mu_\mathrm{b}) = \delta(\mu_\mathrm{b} - \tilde{\mu}_\mathrm{b}). \tag{5}$$

Following the procedure from [8] to write $p(X|N, \mu_\mathrm{s}, \mu_\mathrm{b})$ and marginalizing over the background, we arrive at

$$p(\mu_\mathrm{s}|N, X) \propto \sum_{N_\mathrm{s}=0}^{N} C(X, N_\mathrm{s}) \mu_\mathrm{s}^{N_\mathrm{s}} \mu_\mathrm{b}^{N-N_\mathrm{s}} e^{-(\mu_\mathrm{s}+\mu_\mathrm{b})}. \tag{6}$$

2.1 Implementation of BASiL in Gammapy

Gammapy is an open-source Python package for gamma-ray astronomy in active development, aimed at high-level data analysis [9]. This encompasses data level 3 (DL3), such as event lists and instrument response functions (IRFs), followed by data reduction and modelling to obtain the science results. In this work, new `Dataset` classes were created in a personalized Gammapy version 1.3 [12], so that the BASiL method could be incorporated into the existing Gammapy framework.

Both new `SpectrumDatasetOnOffBASiL` and `MapDatasetBASiL` classes inherit the standard `SpectrumDatasetOnOff` and `MapDataset`, respectively, replacing the computation of the statistic for each bin – the `stat_array()` method – by the implementation of eqs. 3 and 6. In the BASiL case this was defined to be $\lambda \equiv -2 \ln p(\mu_\mathrm{s}| \dots)$, so the expected signal can be obtained from the maximum posterior (i.e. minimization of the statistic).

Some optimizations were made to reduce computational time, due to the more complex likelihood. Firstly, the factor $C(X, N_\mathrm{s})$ alongside all terms without μ_s or μ_b are computed one single time, for all bins chosen in the analysis. This array is then stored to be read during the statistic calculation. Specifically for the 3D analysis, an auxiliary Cython function is used to perform the sum over `stat_array()` – the `stat_sum()` method –, similar to the one already implemented in Gammapy for the Cash statistic, which greatly speeds up computation time. However, while BASiL's 1D `stat_array` is ~5 times slower (CPU time), the factor still increases to ~100 in the 3D analysis.

3. Data analysis

Two types of IRFs were used to sample events and perform the forward folding fitting analysis. For the standard approach, we considered IRFs from the public data release of H.E.S.S. [13], which were created including gamma-hadron separation cuts based on Hillas parametrization. A second set of IRFs was produced for this work destined for BASiL analysis, in which no cuts were applied and a BDT is used for gamma-hadron spearation. With these IRFs, alongside a defined source model, event lists were sampled using Gammapy. To sample the background in the situation of no cuts, the background model was scaled by a factor of 30, which is a value estimated to approximate the amount of background events that are eliminated when a standard cut is applied, although this procedure do not take into account its particular energy dependence. To each event, we attributed a BDT classification value, sampled from the corresponding expected distribution of gamma rays or background. These distributions originated, respectively, from Monte Carlo simulations and a set of Off runs; and are also required to compute $C(X, N_s)$, by evaluating the likelihoods $p(x_i|\gamma)$ and $p(x_i|\bar{\gamma})$.

3.1 1D analysis

The IRFs associated with an observation of PKS 2155-304 at a zenith angle of 36.8 deg were used to sample a point source at 0.5 deg offset with a power-law spectrum, considering a livetime of approximately 26 minutes. Two values of the normalization parameter were chosen (3×10^{-12} and 3×10^{-11} TeV^{-1} cm^{-2} s^{-1}, fixing the reference energy $E_0 = 1$ TeV), corresponding to 7.5% and 75% of the Crab flux at 1 TeV, as well as three spectral indices (1.5, 2 and 3.5). For each combination of parameters, 500 event lists were sampled considering the IRFs with and without cuts. Then, the respective standard and BASiL analyses were performed to obtain the best-fit parameters and compute flux points when possible. Figure 1 shows the relative uncertainty (uncertainty over reconstructed parameter value) for the spectral index parameter, considering the median of the sampled distribution and the 16% and 84% quantiles. A very similar trend is also observed for the amplitude.

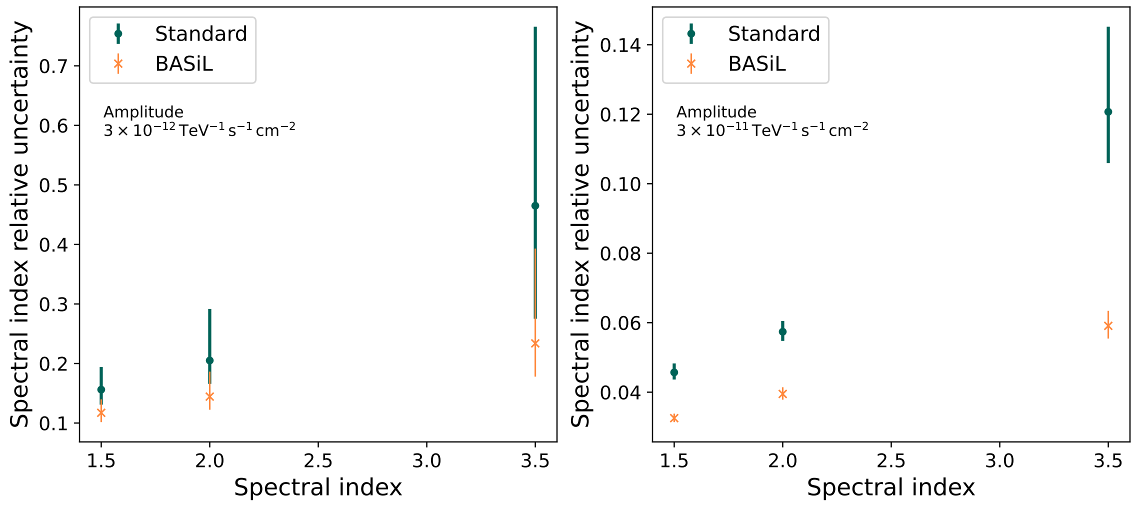

Figure 1: Relative uncertainty of the spectral index (uncertainty over parameter value), as a function of the true spectral index, for two amplitude values ("low" and "high" on the left and right, respectively). The points indicate the median of 500 pseudo experiments, with the error bars corresponding to the 16% and 84% quantiles.

Regardless of the amplitude or spectral index value, the BASiL approach consistently results in lower uncertainties of the reconstructed spectral parameters. The spread of the distribution is also much smaller, which is particularly noticeable for the low amplitude and soft-spectrum source (index 3.5), as the reconstruction of the spectral features in the standard method is particularly affected by the low statistics. Taking the ratio between the BASiL and standard median relative uncertainties, Figure 2 reveals the improvement in uncertainty for both spectral index (left) and amplitude parameters (right). Compared to the standard method, this ranges from around ~25% at spectral index 1.5 to ~50% at index 3.5, showing a stronger impact of the BASiL method to softer spectra in this case. Since the energy threshold is close to the reference energy, there are typically less signal events with softer indices. With brighter sources (higher amplitude), we see compatible ratios between the uncertainties, as well as a noticeable reduction in the spread of the reconstructed parameters.

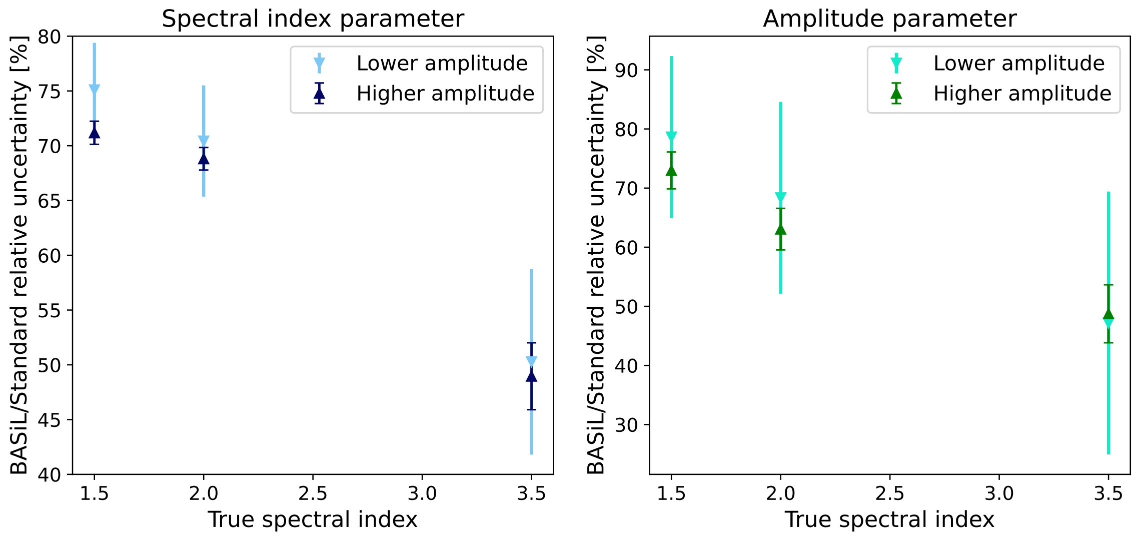

Figure 2: Ratio of the median relative uncertainties between the BASiL and standard methods, with respect to the spectral index (left) and amplitude parameter (right). The upward triangles describe the high amplitude case ($3{\times}10^{-11}$ TeV^{-1} cm^{-2} s^{-1}), while the downward triangles, the low amplitude ($3{\times}10^{-12}$ TeV^{-1} cm^{-2} s^{-1}). The uncertainties are the propagated median absolute deviations of each distribution.

3.2 3D analysis

For the 3D analysis, based on eq. 6, we considered again a point-source with power-law spectrum (spectral index $\Gamma = 2$ with amplitudes 3×10^{-12} and 3×10^{-11} TeV^{-1} s^{-1} cm^{-2}). Since the background template without BDT cuts is not ready, we sampled a source using the standard IRFs available in the public data release of H.E.S.S. [13] and simply scaled the background model by a factor of 30. In this case, we considered a total of four observation runs (each with a livetime of approximately 28 minutes) with a symmetric pointing position 0.5 deg offset in relation to the source, adopting the IRFs corresponding to the observation of PKS 2155-304 at zenith angle of ~50 deg.

Once the datasets were produced and the necessary BASiL factors were computed, a fit of a point-source and power-law models was performed. Figure 3 shows the resulting spectral model

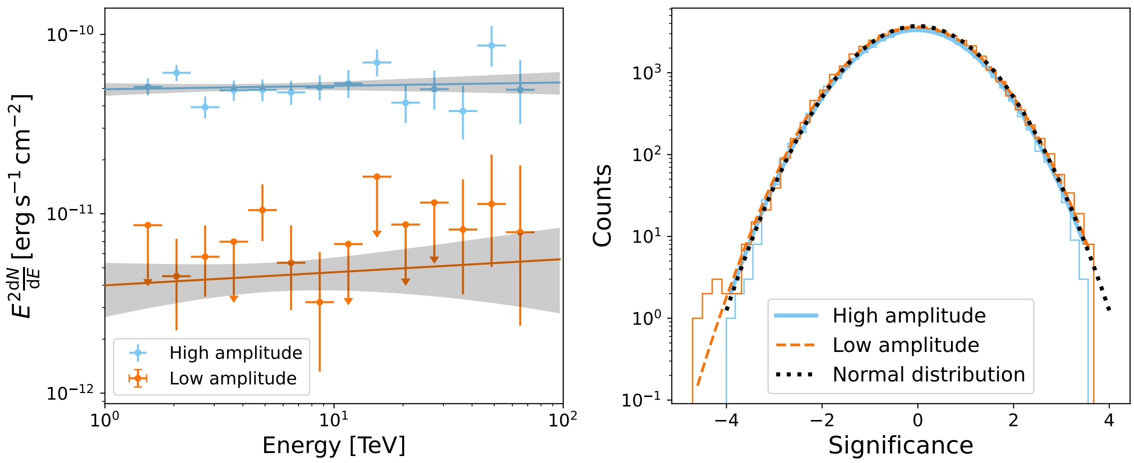

Figure 3: *Left*: Spectral models and flux points from the point source plus power-law model fitting using the 3D BASiL implementation. Low amplitude corresponds to 3×10^{-12} TeV^{-1} s^{-1} cm^{-2} and high amplitude to 3×10^{-11} TeV^{-1} s^{-1} cm^{-2}. *Right*: significance distribution of the dataset maps after modelling, consistent to a normal curve, as expected for an empty map with random fluctuations.

and flux points for both amplitude values, as well as the significance distribution of the map after modelling (which closely matches a normal distribution, as expected). The results support the validity of the 3D BASiL implementation, as the reconstructed parameters agree with the true values – index 1.981 ± 0.045 and amplitude $(3.09 \pm 0.25) \times 10^{-11}$ TeV^{-1} s^{-1} cm^{-2} for the bright source and, for the faint case, the values 1.93 ± 0.16 and $(2.49 \pm 0.83) \times 10^{-12}$ TeV^{-1} s^{-1} cm^{-2} –, while no residual structures are observed in the significance map.

4. Conclusion

In this proceeding, we showed a practical implementation of BASiL in Gammapy that combines the Bayesian approach to signal estimation with the forward folding method, to reconstruct spatial and spectral characteristics from model fitting. By comparing BASiL with the standard procedure, in which gamma-hadron separation cuts are applied, we achieved uncertainty improvements in the range of 20–50%, due to increased signal counts and additional information in the likelihood. We also present an extension of BASiL that allows for combined spatial and spectral analyses, analogous to the standard Cash statistic. Although more computationally intensive, this approach expands the scope of applications, such as providing a tool for faint extended sources. In the future, investigations on systematic errors are expected to be done, along with applications to real observations.

Acknowledgments

The authors thank H.E.S.S. Collaboration for providing the simulated data and common analysis tools that allowed this work. The authors also thank Edivaldo Moura Santos for helping to organize this project. M.G.D.X. acknowledges Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for grant #2022/07692-4.

References

[1] J. Albert, E. Aliu, H. Anderhub, P. Antoranz, A. Armada, M. Asensio et al., *Implementation of the Random Forest method for the Imaging Atmospheric Cherenkov Telescope MAGIC*, *Nuclear Instruments and Methods in Physics Research A* **588** (2008) 424 [0709.3719].

[2] S. Ohm, C. van Eldik and K. Egberts, *γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analysis method*, *Astroparticle Physics* **31** (2009) 383 [0904.1136].

[3] F. Aharonian, A.G. Akhperjanian, A.R. Bazer-Bachi, M. Beilicke, W. Benbow, D. Berge et al., *Observations of the Crab nebula with HESS*, A&A **457** (2006) 899 [astro-ph/0607333].

[4] S. Ohm, S. Wagner and H. E. S. S. Collaboration, *Current status and operation of the H.E.S.S. array of imaging atmospheric Cherenkov telescopes*, *Nuclear Instruments and Methods in Physics Research A* **1055** (2023) 168442 [2306.02716].

[5] T.P. Li and Y.Q. Ma, *Analysis methods for results in gamma-ray astronomy.*, ApJ **272** (1983) 317.

[6] W. Cash, *Parameter estimation in astronomy through application of the likelihood ratio.*, ApJ **228** (1979) 939.

[7] T.J. Loredo, *Promise of bayesian inference for astrophysics*, in *Statistical Challenges in Modern Astronomy*, E.D. Feigelson and G.J. Babu, eds., (New York, NY), pp. 275–297, Springer New York (1992), DOI.

[8] G. D'Amico, T. Terzić, J. Striškovič, M. Doro, M. Strzys and J. van Scherpenberg, *Signal estimation in on/off measurements including event-by-event variables*, Phys. Rev. D **103** (2021) 123001 [2105.01019].

[9] A. Donath, R. Terrier, Q. Remy, A. Sinha, C. Nigro, F. Pintore et al., *Gammapy: A python package for gamma-ray astronomy*, *A&A* **678** (2023) A157.

[10] D. Berge, S. Funk and J. Hinton, *Background modelling in very-high-energy γ-ray astronomy*, A&A **466** (2007) 1219 [astro-ph/0610959].

[11] S.S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Annals Math. Statist.* **9** (1938) 60.

[12] F. Acero, A. Aguasca-Cabot, J. Bernete, N. Biederbeck, J. Djuvsland, A. Donath et al., *Gammapy: Python toolbox for gamma-ray astronomy*, Jan., 2025. 10.5281/zenodo.14760974.

[13] H. E. S. S. Collaboration, *H.E.S.S. first public test data release*, *arXiv e-prints* (2018) arXiv:1810.04516 [1810.04516].