PROCEEDINGS OF SCIENCE

Towards Multi-Instrument and Reproducible Gamma-Ray Analysis

Cosimo Nigro*
Deutsches Elektronen-Sychrotron DESY, Zeuthen, Germany
E-mail: cosimo.nigro@desy.de

Kai Brügge†
Technische Universität Dortmund, Germany
E-mail: kai.bruegge@tu-dortmund.de

The analysis and combination of data from different gamma-ray instruments usually involves the use of a multitude of software packages and data formats. Driven by the requirements for the upcoming Cherenkov Telescope Array (CTA) and a desire for easily reproducible results, an effort has been started to define a common data format for high-level data, namely event lists and instrument response functions (IRFs). In this collaborative project between the FERMI, MAGIC, VERITAS, FACT, and H.E.S.S experiments, a common data format was formalized and implemented. Data from observations of the Crab Nebula was made open to the public and analyzed with the open-source `gammapy` software. By combining data from FERMI, and the four currently operating imaging atmospheric Cherenkov telescopes, we produced a joint maximum likelihood fit of the Crab Nebula spectrum. Aspects of the statistical errors and the evaluation of systematic uncertainties are commented upon. All datasets and results presented in this work can be obtained using open-access on-line assets that allow for long-term reproducibility for everyone.

36th International Cosmic Ray Conference -ICRC2019-
July 24th - August 1st, 2019
Madison, WI, U.S.A.

*Thanks to all authors who do not fit on this title page : C. Deil ,R. Zanin ,T. Hassan ,J. King ,J.E. Ruiz ,L. Saha ,R. Terrier ,M. Nöthe ,R. Bird ,T. T. Y. Lin ,J. Aleksić ,C. Boisson ,J.L. Contreras ,A. Donath ,L. Jouvin ,N. Kelley-Hoskins ,B. Khelifi ,K. Kosack ,J. Rico ,A. Sinha

†Speaker.

PoS(ICRC2019)641

1. Introduction and Motivation

This proceeding summarizes the most important results of our previously published paper entitled *Towards open and reproducible multi-instrument analysis in gamma-ray astronomy* [1]. Combining images from multiple astronomical facilities is key to unravel some of the mysteries of modern astrophysics. In the late 1970s, in an effort to overcome the technical challenges involved when sharing images among operating systems, a common astronomical format was developed. The Flexible Image Transport system (FITS) was first standardized in 1981 [2]. Today the FITS format is in widespread use among astronomers of all observing bands, from radio frequencies to gamma rays. The HE gamma-ray (E>100 MeV) Large Area Telescope [3], on board the *Fermi* satellite, publicly releases all its high-level analysis data in FITS format. *Fermi*'s analysis software, the science tools, are publicly available as well and allow anyone to compute spectra, light curves, and sky maps. The astroparticle community, and very-high-energy (VHE;E>100 GeV) gamma-ray astronomy in particular, inherited its methodologies and traditions from particle physics. Ground-based observation of very-high-energy gamma-rays is performed by Imaging Atmospheric Cherenkov Telescopes (IACTs) that record event-based data instead of images. All currently operating IACTs rely on the `ROOT` [4] framework and its file format which is so ubiquitous in particle physics. Despite the common container format neither the internal data structure nor the software is shared among the different experiments. Our paper aims to show that the FITS data format can accommodate gamma-ray data and that scientific results can be obtained with python-based open-source software. In a joint effort, we collected data from the H.E.S.S. [5], VERITAS [6], H.E.S.S. [5], and FACT [7] telescopes and published event data in a common data format including the corresponding instrument response functions. We perform a spectral analysis using the combined data from each IACT together with *Fermi*-LAT data using the open-source `gammapy` [8] science tool.

2. Datasets

Unlike typical astronomical telescopes, instruments for gamma-ray astronomy cannot directly scatter or reflect gamma rays. The experimental techniques, either space-borne or ground-based [9], rely on the direct detection of secondary charged particles or on the indirect detection of the Cherenkov emission of a cascade of charged secondaries they produce in the atmosphere. A detection, or event, cannot unambiguously be discriminated from the irreducible charged cosmic-ray background, but can only be classified with a certain probability as a primary photon. In a similar manner, the energy and direction of the primary photon can only be estimated from the secondary products. The input for the high-level analysis of gamma-ray astronomy data typically consists of two parts. First, a list of events that are classified as gamma-ray candidates along with their estimated direction, estimated energy, and arrival time. Second, the instrument response functions (IRFs) that quantify the performance of the detector and connect the estimated quantities with their true, physical, values. As the true values are unknown during observation, the IRFs have to be computed from exhaustive simulations. For the joint analysis presented here, all participating IACTs released event-lists and IRFs for observations of the Crab Nebula. The Crab Nebula is the reference source of VHE gamma-ray astronomy due to its brightness and flux steadiness. Table 1 summarizes the individual data sets. The published data was converted into the FITS-based com-

PoS(ICRC2019)641

Table 1: The datasets were selected to contain approximately equal amounts of counts. Due to the different collection areas of the instruments, the observation times differ between instruments. FACT, being the smallest IACT, requires 10 hours of observation whereas MAGIC and VERITAS only need 40 minutes.

Dataset	Year	Duration	Mode	Energy Range / TeV
Fermi-LAT		8 years	Sky Survey	0.03 – 2
MAGIC	2013	40 minutes	Pointing	0.03 – 2
VERITAS	2011	40 minutes	Pointing	0.15 – 30
H.E.S.S.	2004	3 hours	Pointing	0.5 – 30
FACT	2013	10 hours	Pointing	0.4 – 30

mon data format. The format was initiated within the "Data formats for gamma-ray astronomy" forum [10]. This endeavour brings together members of different IACT collaborations with a focus on the upcoming Cherenkov Telescope Array (CTA). Open standards are of utmost importance to CTA since it will be the first VHE gamma-ray experiment that is operated as an open observatory. Figure 1 shows the observed event counts recorded by each individual instrument.

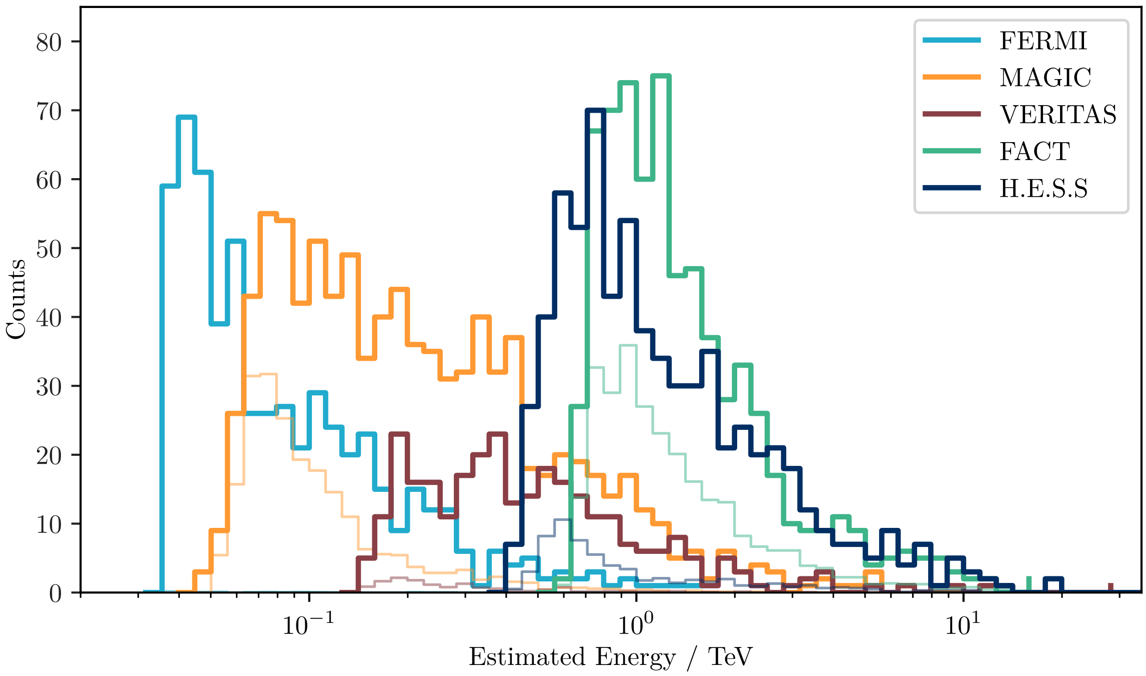

Figure 1: The figure shows the of counts recorded by each individual instrument with respect to estimated energy. The thick lines indicate the number of counts in the signal region while the thin lines represent the number of background counts.

3. Spectral Fit

We perform a fit of the Crab Nebula's energy spectrum for each individual instrument as well as the joint data. Signal and background counts are extracted from the event-list using aperture photometry techniques. The irreducible background in IACT data is approximated by collecting events that originate in region devoid of gamma-ray emitters in the sky, the OFF region. From the even-lists we extract the total number of events in the OFF region, N_{off}, and the total number of events in the signal, or ON, region N_{on}. The hypothetical counts from the source N_{sig} can then be modeled as

$$N_{\text{on}} = N_{\text{sig}} + \alpha N_{\text{off}} \tag{3.1}$$

where α is the exposure ratio between the ON and OFF regions. Let μ_s be the expected number of signal counts from the source and $\alpha\mu_b$ the expected background counts, the counts in the ON and OFF regions are then distributed as

$$P_{\text{on}} \sim \text{Poisson}\left(\mu_s + \alpha\mu_b\right) \quad \text{and} \quad P_{\text{off}} \sim \text{Poisson}\left(\mu_b\right). \tag{3.2}$$

From this we build a likelihood for the counts in each energy bin i by multiplying the distributions for the signal and the background

$$\mathscr{L}\left(N_{\text{on}}, N_{\text{off}}, \alpha \mid \mu_s, \mu_b\right) = \prod_i P_{\text{on},i}\left(N_{\text{on},i} \mid \mu_{s,i} + \alpha\mu_{b,i}\right) \cdot P_{\text{off},i}\left(N_{\text{off},i} \mid \mu_{b,i}\right). \tag{3.3}$$

The number of expected counts in the signal region μ_s is predicted by folding the assumed spectral model with the IRFs of each instrument. We assume a log-parabolic spectral model for the inverse compton emission of the Crab Nebula as suggested by previous studies [11].

$$\frac{\mathrm{d}\phi}{\mathrm{d}E} = \phi_0 \left(\frac{E}{E_0}\right)^{-\Gamma - \beta \log_{10}\left(\frac{E}{E_0}\right)}. \tag{3.4}$$

The spectral parameters (ϕ_0, Γ, β) are left free to vary in the fit while the reference energy E_0 is fixed at the value of 1 TeV. No model for the background is included in the data. Hence, we consider the μ_b to be nuisance parameters and fix them to the value returning $\frac{\partial \log(\mathscr{L})}{\partial \mu_{b,i}}$. The full likelihood is minimized using `gammapy`'s interface to the `minuit` [12] minimization tool. The resulting spectral energy distributions are shown in figure 2, together with a theoretical model taken from [13]. For a full table of results we refer to our paper [1].

4. Systematic Uncertainties

As mentioned, the energy of the primary photon can only be estimated from its secondary products. The reconstruction of the energy usually employs machine learning methods which are trained on instrument specific simulations. To account for the systematic uncertainties on the energy scale of the different instruments while performing the joint likelihood fit, we use a method shown by H. Dembinski et. al. in [14]. We introduce a nuisance parameter, $z_i = \frac{\tilde{E}-E}{E} = \frac{\tilde{E}}{E} - 1$, relating the reported ($\tilde{E}$) and true (E) energy per instrument. The differential energy flux model is modified to:

$$\frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}\tilde{E}} = \phi_0 \left(\frac{E/(1+z)}{E_0}\right)^{-\Gamma + \beta \log_{10}\left(\frac{E/(1+z)}{E_0}\right)} \left(\frac{1}{1+z}\right) \tag{4.1}$$

The global likelihood function is extended with the distributions of the nuisance parameters z. It is assumed to follow a Gaussian distribution with mean 0 and variance σ_i^2, i.e. the uncertainty in the energy reconstruction estimated by each instrument. Figure 3 shows the estimated error contours for the fits including the systematic uncertainty.

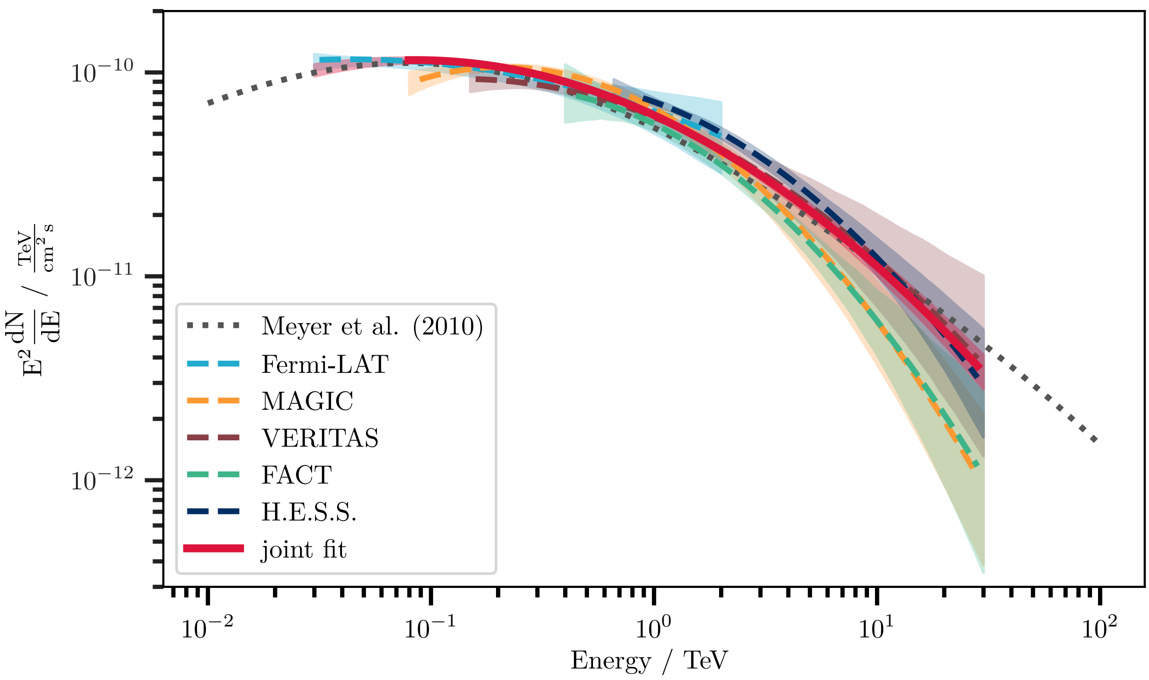

Figure 2: Resulting Crab Nebula spectral energy distribution (SED) from individual instruments and from the joint fit. Only statistical error bands are displayed. The gray line indicates the Crab Nebula SSC model as fitted by Meyer et. al. [13]. The error bands are built from sampling multiple solutions from the covariance matrix which resulted from the `minuit` fit. The statistical power of the joint fit is evident by the smaller error band surrounding the red line.

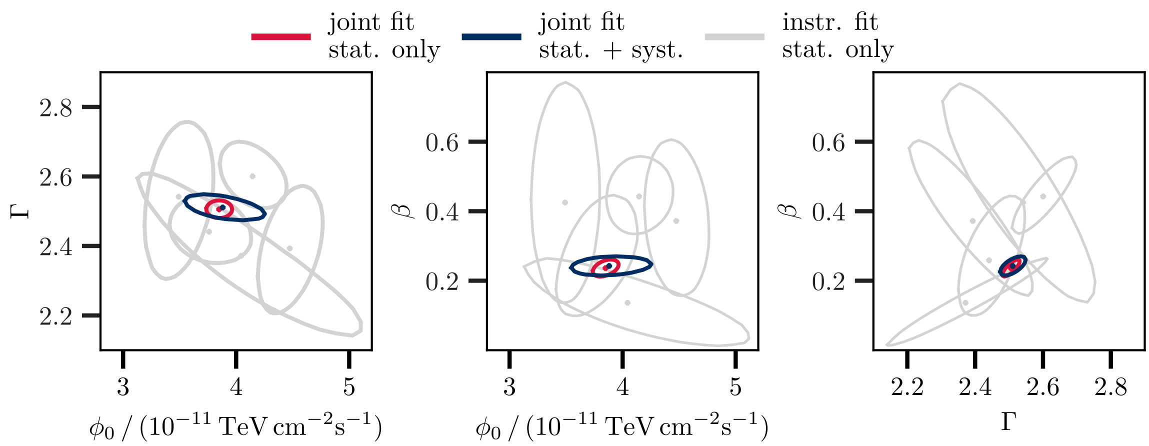

Figure 3: The figure shows the likelihood contours for the joint fit. The red contour only shows statistical errors while the blue line models systematic effects. As expected, the additional nuisance parameter increases the uncertainty in th estimated parameters.

5. Conclusions

We have shown that the standardized DL3 format allows for the first time to run a multi-instrument analysis using data from *Fermi*-LAT and all the existing IACTs. This format, proposed for the future Cherenkov Telescope Array (CTA), allows reproducible science in the very-high-energy (VHE) field relying on open-source software. All datasets, scripts and Jupyter notebooks are provided in a GitHub repository together with a Docker image published in DockerHub. This will allow any user to reproduce the results in an interactive environment. For more details we refer the reader to our full paper [1] and to our online repository

https://github.com/open-gamma-ray-astro/joint-crab

References

[1] C. Nigro, C. Deil, R. Zanin, T. Hassan, J. King, J. E. Ruiz, L. Saha, R. Terrier, K. Brügge, M. Nöthe, R. Bird, T. T. Y. Lin, J. Aleksić, C. Boisson, J. L. Contreras, A. Donath, L. Jouvin, N. Kelley-Hoskins, B. Khelifi, K. Kosack, J. Rico, and A. Sinha, *Astronomy & Astrophysics* **625** (May, 2019).

[2] D. C. Wells, E. W. Greisen, and R. H. Harten, *Astronomy & Astrophysics Supplement* **44** (June, 1981) 363.

[3] **Fermi** Collaboration, W. B. Atwood et al., *Astrophysical Journal* **697** (6, 2009) 1071–1102.

[4] R. Brun and F. Rademakers, *Nuclear Instruments and Methods in Physics Research A* **389** (Feb, 1997) 81–86.

[5] J. A. Hinton et al., *New Astronomy Review* **48** (Apr., 2004) 331–337.

[6] J. Holder et al., *Astroparticle Physics* **25** (July, 2006) 391–401.

[7] H. Anderhub et al., *Journal of Instrumentation* **8** (2013) P06008.

[8] C. Deil, R. Zanin, J. Lefaucheur, C. Boisson, B. Khelifi, R. Terrier, M. Wood, L. Mohrmann, N. Chakraborty, J. Watson, R. Lopez-Coto, S. Klepser, M. Cerruti, J. P. Lenain, F. Acero, A. Djannati-Ataï, S. Pita, Z. Bosnjak, C. Trichard, T. Vuillaume, A. Donath, J. King, L. Jouvin, E. Owen, B. Sipocz, D. Lennarz, A. Voruganti, M. Spir-Jacob, J. E. Ruiz, and M. P. Arribas, *International Cosmic Ray Conference* (Jan., 2017).

[9] S. Funk, *Annual Review of Nuclear and Particle Science* **65** (Oct., 2015) 245–277.

[10] C. Deil, C. Boisson, K. Kosack, J. Perkins, J. King, P. Eger, M. Mayer, M. Wood, V. Zabalza, J. Knödlseder, T. Hassan, L. Mohrmann, A. Ziegler, B. Khelifi, D. Dorner, G. Maier, G. Pedaletti, J. Rosado, J. L. Contreras, J. Lefaucheur, K. Brügge, M. Servillat, R. Terrier, R. Walter, and S. Lombardi, *Open high-level data formats and software for gamma-ray astronomy*, in *6th International Symposium on High Energy Gamma-Ray Astronomy*, vol. 1792 of *American Institute of Physics Conference Series*, Jan., 2017.

[11] **MAGIC** Collaboration, J. Aleksić et al., *Journal of High Energy Astrophysics* **5-6** (2015) 30 – 38.

[12] F. James and M. Roos, *Computer Physics Communication.* **10** (1975).

[13] M. Meyer et al., *Astronomy & Astrophysics* **523** (Nov., 2010).

[14] H. Dembinski et al., *Data-driven model of the cosmic-ray flux and mass composition from* 10 *GeV to* 10^{11} *GeV*, in *Proceedings, 35th International Cosmic Ray Conference (ICRC2017): Busan, South Korea, July 2017*, vol. 35, Jan., 2017.