The joint-crab project: towards open and reproducible multi-instrument analysis in gamma-ray astronomy

Gammapy General User Call

C. Nigro [cosimo.nigro@ifae.es] on behalf of the authors in A&A 625, A10 (2019)

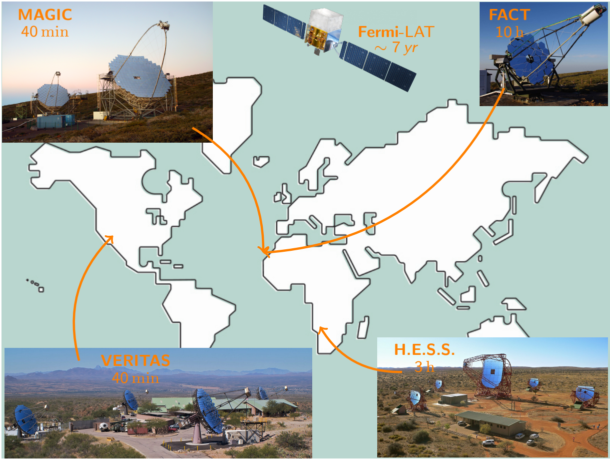
26 October 2020

Context

- Neither data format nor software shared by gamma-ray instruments;
- current perspective, two setbacks:
  - → combination of data from different experiments requires custom expansions of proprietary software;
  - → release of public legacy data needs release of analysis software;
- forward perspective, a challenge:
  - → operation of CTA as an observatory poses VHE community the problem of producing public data and analysis tools.
- Effort started with the *Data Formats for Gamma-ray Astronomy* forum to define a (prototypical) standardised format for high-level gamma-ray data (DL3 - see Regis' talk).

The joint-crab Project

- Using this preliminary DL3 data format, we performed the **first fully-reproducible multi-instrument gamma-ray analysis**;
- relying on open-source software: `gammapy`;
- combining data from Fermi-LAT, and four of the currently-operating IACTs, to produce **a joint fit of the Crab Nebula spectrum**;
- online material (data and scripts) at: https://github.com/open-gamma-ray-astro/joint-crab/;
- DISCLAIMER: the purpose of this project is to show a method, not to provide a new measurement of the Crab Nebula spectrum.



- *Fermi*-LAT data freely available, IRF computed with the *Fermi* science tools and made DL3-compliant with `gammapy`;
- small samples of DL3 data released specifically for this project†, **first joint release of data by IACT collaborations**.

† H.E.S.S. and FACT data already publicly available

> Event List

Current Table Properties

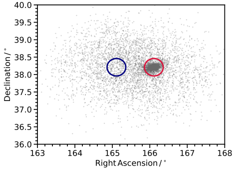
Label: hess_dl3_dr1_obs_id_020136.fits.gz
Location: /Users/cosimo/work/gammapy-data/hess-dl3-dr1/data/hess_dl3_dr1_ol
Name: EVENTS-1
Rows: 11,243
Columns: 5

TOPCAT(1): Table Browser

Table Browser for 1: hess_dl3_dr1_obs_id_020136.fits.gz

	EVENT_ID	TIME	RA	DEC	ENERGY
1	1808181231761	1.019626E8	229.239	-58.3417	0.558903
2	1808181231778	1.019626E8	226.886	-58.921	0.671012
3	1808181231793	1.019626E8	230.858	-57.9829	5.21256
4	1808181231871	1.019626E8	229.745	-60.3202	0.550593
5	1808181231873	1.019626E8	227.885	-60.6502	0.4242
6	1808181231920	1.019626E8	230.813	-59.0734	0.527166
7	1808181231987	1.019626E8	230.767	-57.8558	0.82928
8	1808181231989	1.019626E8	225.231	-60.7951	1.96649
9	1808181232037	1.019626E8	232.746	-58.8183	0.997396

> background estimation ⇒ N_{on} and N_{off} counts



> Instrument Response Function

Current Table Properties

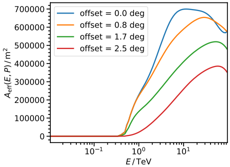
Label: hess_dl3_dr1_obs_id_020136.fits.gz-3
Location: /Users/cosimo/work/gammapy-data/hess-dl3-dr1/data/hess_dl3_dr1_ol
Name: AEFF
Rows: 1
Columns: 5

TOPCAT(3): Table Browser

Table Browser for 3: hess_dl3_dr1_obs_id_020136.fits.gz-3

	ENERG_LO	ENERG_HI	THETA_LO	THETA_HI	EFFAREA
1	(0.01, 0.01...	(0.011006...	(0.0, 0.5, 1...	(0.0, 0.5...	((0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0

Total: 1 Visible: 1 Selected: 0



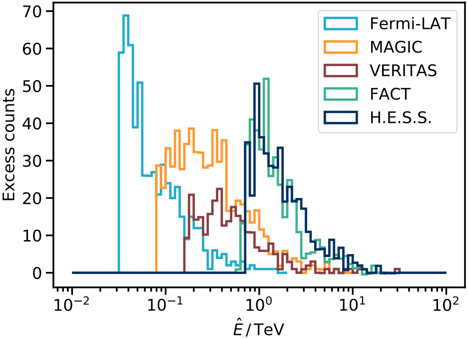
> compute the expected counts

$$g(\mathbf{\Lambda}) = \frac{\mathrm{d}\phi}{\mathrm{dE}}(\mathbf{\Lambda}) * \mathrm{IRF}$$

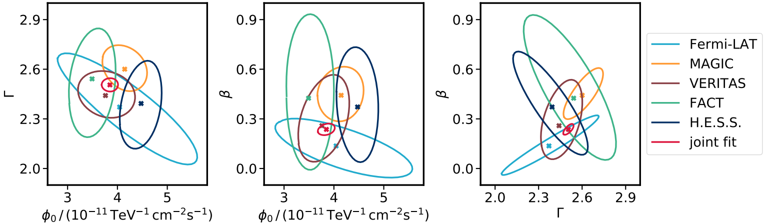
Binned (energy, $n_{\hat{E}}$) Maximum Likelihood method, estimate $\frac{\mathrm{d}\phi}{\mathrm{dE}}(\mathbf{\Lambda})$:

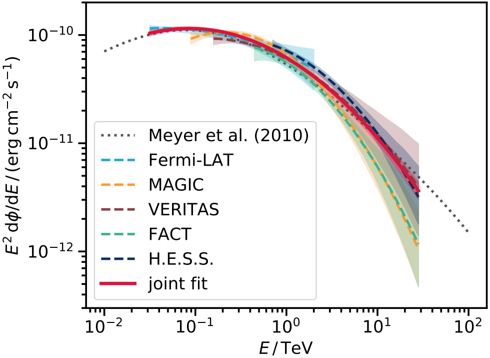
$$\mathcal{L}(\mathbf{\Lambda}|\mathcal{D}) = \prod_{i=1}^{n_{\mathrm{instr}}} \underbrace{\mathcal{L}_i(\mathbf{\Lambda}|\{N_{\mathrm{on},ijk}, N_{\mathrm{off},ijk}\}_{j=1,\ldots,n_{\mathrm{runs}};k=1,\ldots,n_{\hat{E}}})}_{\prod_{j=1}^{n_{\mathrm{runs}}} \prod_{k=1}^{n_{\hat{E}}} \mathrm{Pois}(g_{ijk}(\mathbf{\Lambda})+b_{ijk};N_{\mathrm{on},ijk}) \times \mathrm{Pois}(b_{ijk}/\alpha_{ij};N_{\mathrm{off},ijk})}$$

Counts Spectra



Dataset	T_{obs}	$E_{\mathrm{min}}/\mathrm{TeV}$	$E_{\mathrm{max}}/\mathrm{TeV}$	$N_{\mathrm{on},i}$	$N_{\mathrm{off},i}/\alpha$	$\theta_{\mathrm{on}}/\mathrm{deg}$
Fermi-LAT	$\sim$ 7 yr	0.03	2	578	1.2	0.3
MAGIC	40 mins	0.08	30	784	129.9	0.14
VERITAS	40 mins	0.15	30	289	13.7	0.10
FACT	10 hours	0.45	30	691	272.8	0.17
H.E.S.S.	2 hours	0.71	30	459	27.5	0.11





- Spectrum Fit

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

- **First Crab spectrum** combining data from *Fermi*-LAT and **four of the currently-operating Cherenkov telescopes**.

Modified Likelihood for Systematic Uncertainties

> Example of how to include the systematic uncertainties on the energy scale of the different instruments (following Dembinski et al. 2017):

→ constant energy bias per instrument $z_i = \frac{\tilde{E}-E}{E} = \frac{\tilde{E}}{E} - 1$;

→ modified assumed spectrum $\tilde{E}$:

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

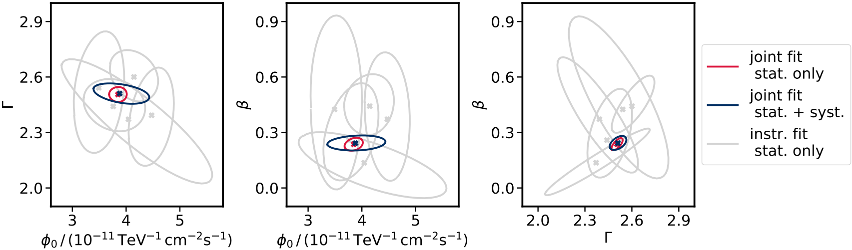
> global likelihood function extended with the distributions of the z_i:

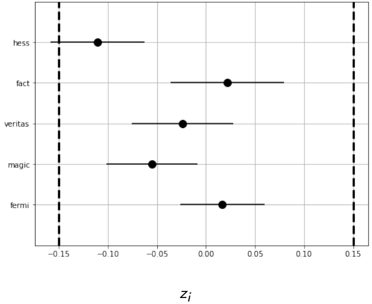
$$\mathcal{L}(\boldsymbol{\Lambda}|\boldsymbol{\mathcal{D}}) = \prod_{i=1}^{n_{\mathrm{instr}}} \mathcal{L}_i(\boldsymbol{\Lambda}|\boldsymbol{\mathcal{D}}_{\mathrm{i}}) \times \mathcal{N}(z_i; 0, \delta_i^2);$$

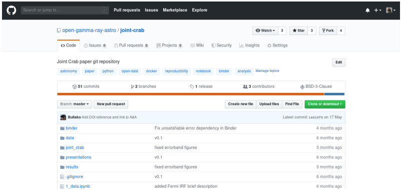
→ z_i fitted with the other spectral parameters;

→ constrained with δ_i = systematic uncertainty on the energy scale estimated by each instrument.

Modified Likelihood for Systematic Uncertainties







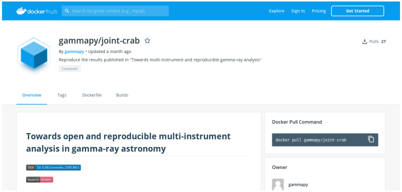
> Short-term:

→ all the code publicly available in `github`;

→ data (∼ MB) shipped with the code;

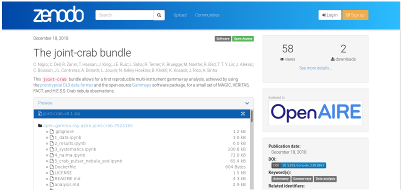
→ simple commands & notebooks to reproduce results;

→ packages managed via `anaconda` environment.



> Medium-term:

→ may happen that conda virtual environment is not enough to guarantee reproducibility (software not anymore mantained), a joint-crab Docker container is provided.



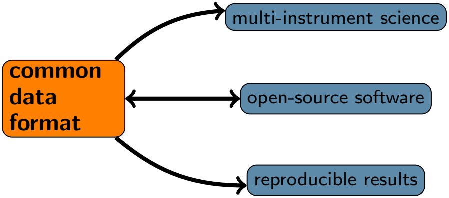
> Long-term:

→ on-line material archived in Zenodo, `DOI: 10.5281/zenodo.2381863`.

Conclusion

- A technical note:
  - → scripts and notebooks for the `joint-crab` project were produced with `gammapy v0.9` (conda and docker to re-run them);
  - → the likelihood fit and the modified likelihood for the syst. estimation have been inserted in the `gammapy-benchmarks`: https://github.com/gammapy/gammapy-benchmarks/tree/master/validation/joint-crab

- **Multi-instrument, reproducible** γ-ray analysis using **open-source software** are already within our reach



- a key asset for future instruments (like CTA) that plan to extend their data access to the wide astrophysical community.