Towards multi-instrument and reproducible gamma-ray analysis

C. Nigro* [1] C. Deil [2] R. Zanin [2] T. Hassan [1,3] J. King [2]
J.E. Ruiz [4] L. Saha [5] R. Terrier [6] K. Bruegge [7] M. Noethe [7]
R. Bird [8] T. T. Y. Lin [9]

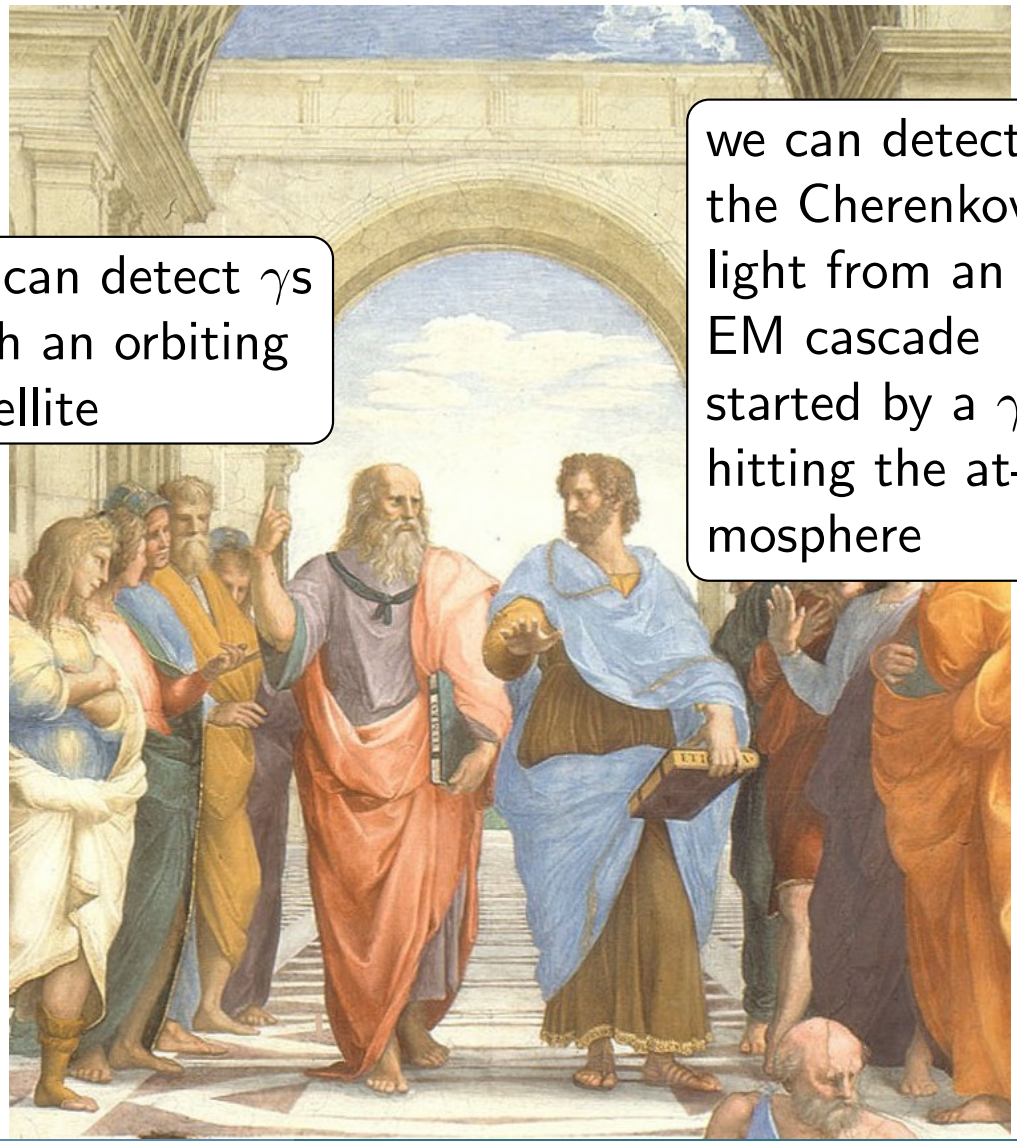
*contact:cosimo.nigro@desy.de
[1]DESY Zeuthen, [2]MPIK Heidelberg, [3]IFAE-BIST Barcelona, [4]IAA-CSIC Granada,
[5]UCM Madrid, [6]APC Paris, [7]TU Dortmund, [8]UCLA, [9]McGill Montreal

TeV Particle Astrophysics, 27-31 August, Berlin, Germany

Introduction

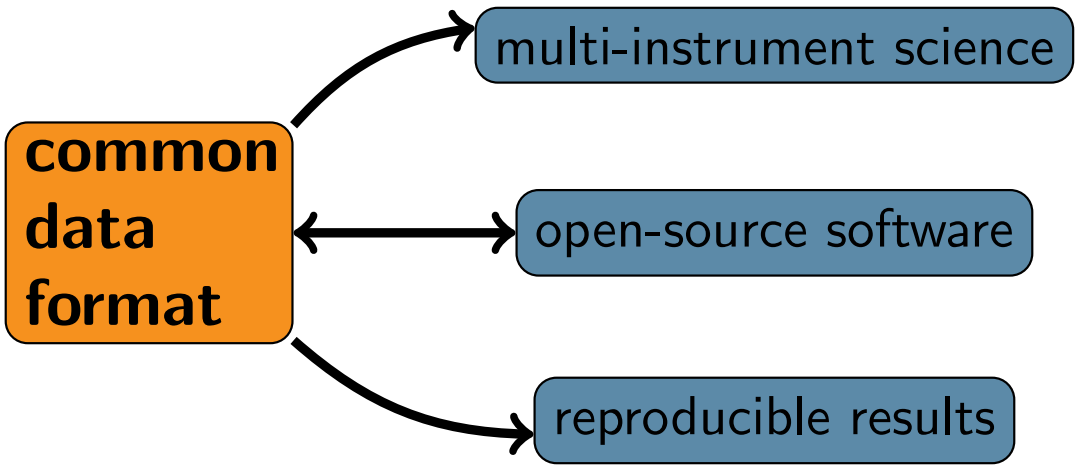
The importance of a common language / format

> Different philosophies, same language.



The importance of a common language / format

- Analysis and combination of data from different gamma-ray instruments today: proprietary software, case-by-case methods - translating Plato to Aristotle!;
- what do we gain by defining a common format for gamma-ray data? - let Plato and Aristotle speak Greek!



A common gamma-ray format

- Community effort already started at *Data formats for gamma-ray astronomy* forum http://gamma-astro-data-formats.readthedocs.io;
- which level to unify?

IACT data level	description	reduction factor
DL0	**raw** output of DAQ	
DL1	**calibrated** quantities (charge, arrival time)	1 - 0.2
DL2	**reconstructed** shower parameters	10^{-1}
DL3	**reduced** γ ray candidates + IRFs	10^{-2}
DL4	**science** data products: spectra, LC, skymaps	10^{-3}
DL5	**observatory** data: surveys, catalogues	10^{-3} - 10^{-5}

- space-borne instrument data (e.g. *Fermi*-LAT) can be embedded in this scheme;
- files stored in FITS format (a 30-year standard in astronomy).

The joint-crab effort

Objectives

- Using this preliminary DL3 format, we perform the first fully-reproducible multi-instrument gamma-ray analysis;
- relying on open-source software: *gammapy*;
- combining data from *Fermi*-LAT, and the four existing IACTs, to produce a joint fit of the Crab Nebula spectrum;
- DISCLAIMER: the purpose of this project is to show a method, not to provide a new measurement of the Crab Nebula spectrum.

Datasets

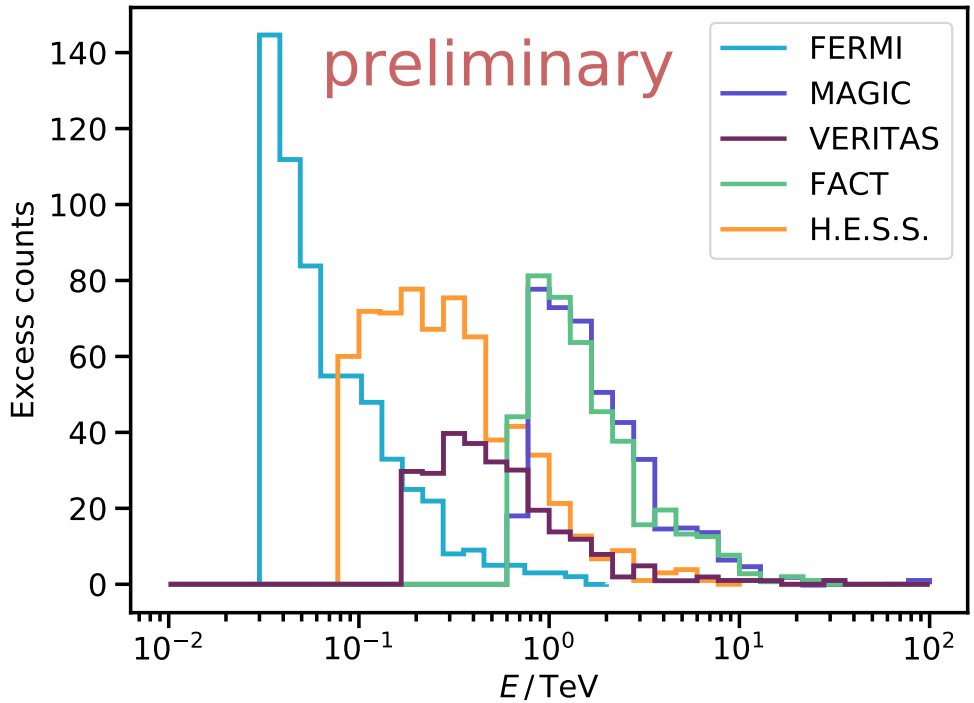
- *Fermi-LAT* data freely available;
- small samples of DL3 data released by IACT collaboration for this project, FACT[1] and H.E.S.S.[2] datasets already available to the public.

Dataset	time	obs. mode	E_{min} / TeV	E_{max} / TeV
Fermi-LAT	$\sim$ 7 years	sky survey	0.03	2
MAGIC	40 mins	pointing	0.08	30
VERITAS	40 mins	pointing	0.15	30
FACT	10 hours	pointing	0.40	30
H.E.S.S.	3 hours	pointing	0.50	30

[1] https://fact-project.org/data/
[2] https://www.mpi-hd.mpg.de/hfm/HESS/pages/dl3-dr1/

Analysis: data reduction

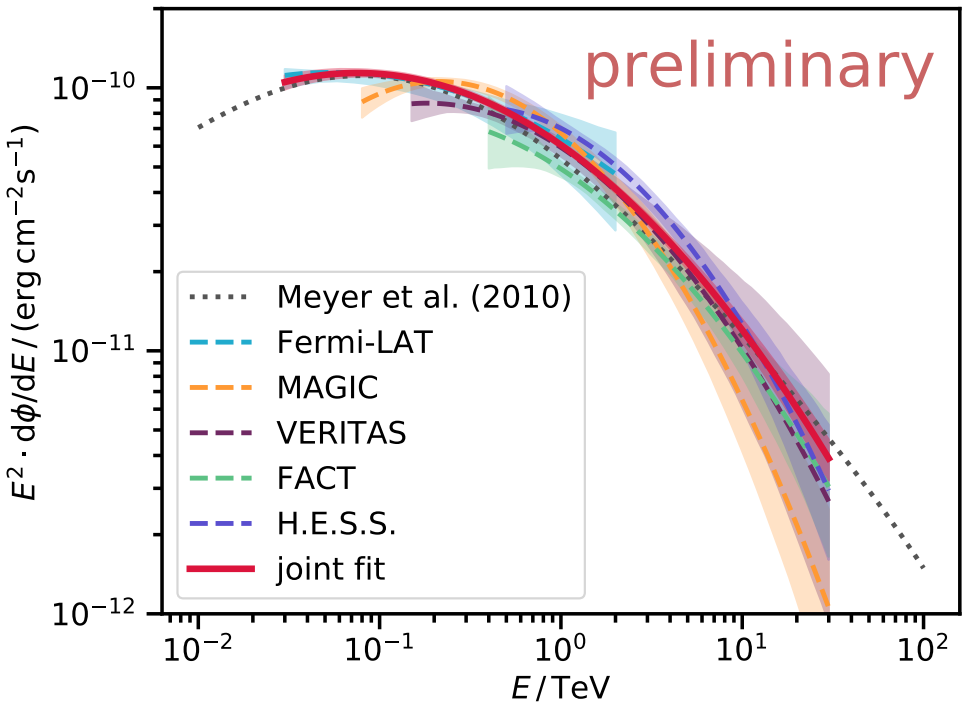
- One-dimensional (energy dependent) spectral likelihood fit:
  - → observed counts: via aperture photometry techniques;



  - → expected counts: folding IRFs with assumed spectral model:

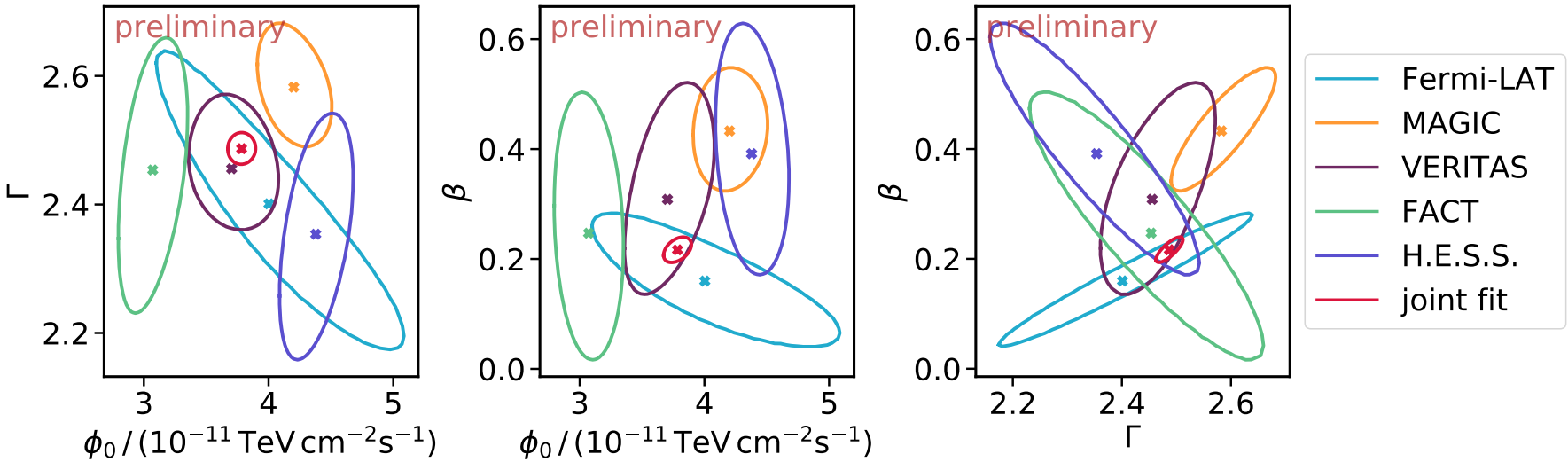
    log-parabola $\frac{d\phi}{dE} = \phi_0 \left(\frac{E}{E_0}\right)^{-\Gamma+\beta \log_{10}\left(\frac{E}{E_0}\right)}$.

Analysis: likelihood fit



> Resulting Crab Nebula SED from individual instruments and from the joint fit.

Analysis: likelihood fit



> Likelihood 1-σ contours for the log-parabola parameters for individual instruments and the joint fit.

Systematics

- Systematic uncertainties on the energy scale of the different instruments accounted for by introducing nuisance parameters:
  - $z = \frac{\tilde{E}-E}{E} = \frac{\tilde{E}}{E} - 1$;
  - differential flux in reconstructed energy $\tilde{E}$

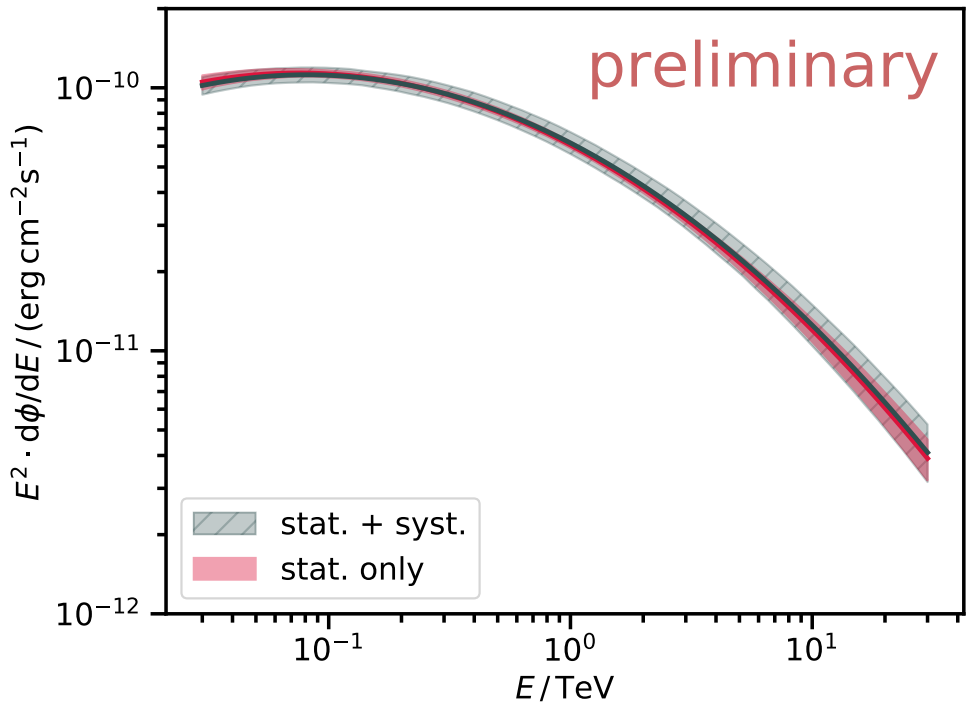
$$\frac{d\tilde{\phi}}{d\tilde{E}} = \phi_0 \left(\frac{E/(1+z_{\rm instr})}{E_0} \right)^{-\Gamma+\beta \log_{10}\left(\frac{E/(1+z_{\rm instr})}{E_0}\right)} \times \left(\frac{1}{1+z_{\rm instr}} \right);$$

- global likelihood function extended with the distributions of the nuisance parameters $z_{\rm instr}$ per each dataset

$$-2 \sum_{\text{all instruments}} \ln \mathcal{L}(\phi_0, \Gamma, \beta, z_{\rm instr} | N_{\rm ON\ instr}, N_{\rm OFF\ instr}) + \left(\frac{z_{\rm instr}}{\delta_{\rm instr}} \right)^2$$

where δ_i = the uncertainty in the energy reconstruction estimated by each instrument.

Systematics



> Resulting Crab Nebula SED from the joint fit including both statistical and systematic error bands.

Using a theoretical model

- An analytical function is not the only possibility to perform a likelihood fit, any theoretical model can be used for $\frac{d\phi}{dE}$.
- Typically theoretical models are not plugged in the likelihood estimation but fitted to spectral points:
  - → often not unfolded (i.e. in $E_{\rm est}$), and limited in cases where the energy dispersion plays a major role.
- Releasing the results of the data reduction (i.e. excess distributions and IRFs) would allow:
  - → successive likelihood fit with any arbitrary theoretical model (example with `naima` in the on-line material of the publication in progress);
  - → later combination with any other MWL data for highly collaborative and extensible future work.

How is reproducibility achieved?

- Short-term:
  - → all the code will be publicly available in `GitHub` https://github.com/open-gamma-ray-astro/joint-crab;
  - → the size of the data is $\sim$ MB, can be provided along with the code;
  - → packages managed via `anaconda` environment.
- Medium-term:
  - → it may happen that the `conda` virtual environment is not enough to guarantee reproducibility (software not anymore mantained), a Docker container will be provided on DockerHub.
- Long-term:
  - → on-line material available on `Zenodo`.

Wrap-up and prospects

An open gamma-ray science

- What can our community achieve?
- An approach to gamma-ray science, summarized by three essential concepts: common data-format, open-source software and fully-reproducible results, the first being the cornerstone of the last two.
- With the `joint-crab` example we illustrate **this approach is already within our reach**: relying on a prototypical DL3 format we combine data from *Fermi*-LAT and the four existing IACTs and make the analysis reproducible within the context of open resources (software, hosting platforms).
- A key asset for future gamma-ray instruments like CTA that will be operated as an open observatory and share its data with a wide astronomical community.