MCPLOTS : MC validation resource based on volunteer computing

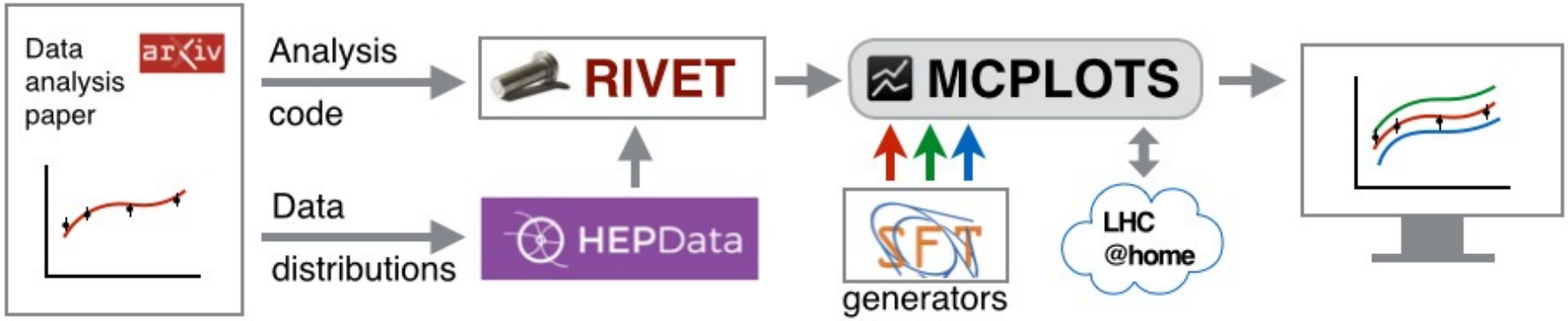
A. Karneyeu[1], N. Korneeva[2,3], P. Skands[2]

[1] Institute for Nuclear Research, Russian Academy of Sciences
[2] Monash University
[3] National Research Tomsk Polytechnic University

Workshop on the tuning of hadronic interaction models
24 January 2024

MCPLOTS : overview

Online repository of Monte Carlo plots compared to experimental data



http://mcplots.cern.ch/

https://rivet.hepforge.org/ https://ep-dep-sft.web.cern.ch/

https://www.hepdata.net/ https://lhcathome.web.cern.ch/

MCPLOTS

- prepare configuration files ;
- run Rivet codes with different generators ;
- collect obtained distributions ;
- provide access to the collection via the website

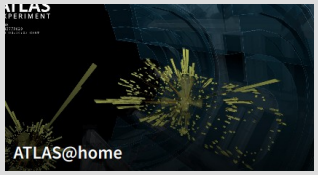
MCPLOTS and LHC@home

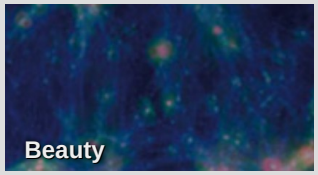
LHC @home — Volunteer computing for the LHC

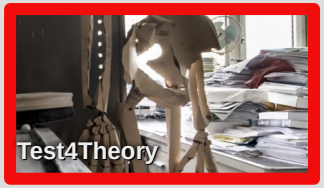
Open Eng. 7 (2017) 1, 378-392

computing platform where voluteers donate idle time on their computers

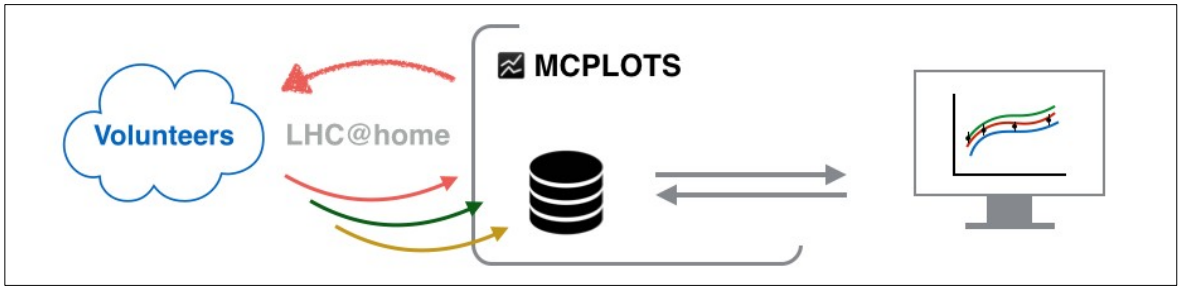
5 projects :







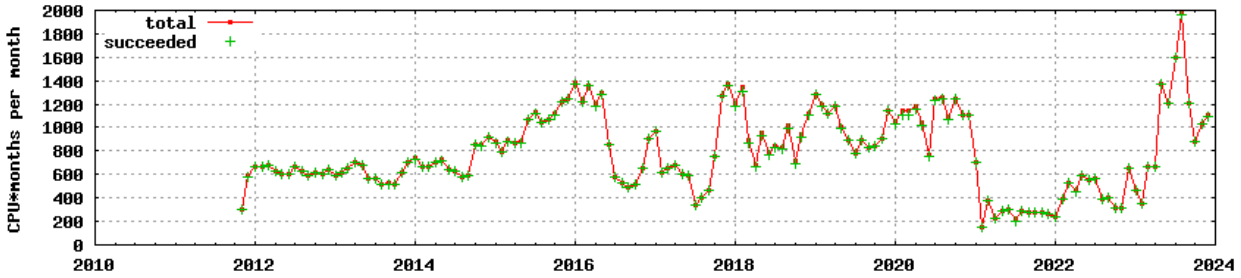
PoS ISGC2012 (2012) 036



MCPLOTS workflow

Jobs are distributed to volunteers
Completed ones are stored on the **server**
Their descriptions – in the **database**
The website operates with **queries** to this DB

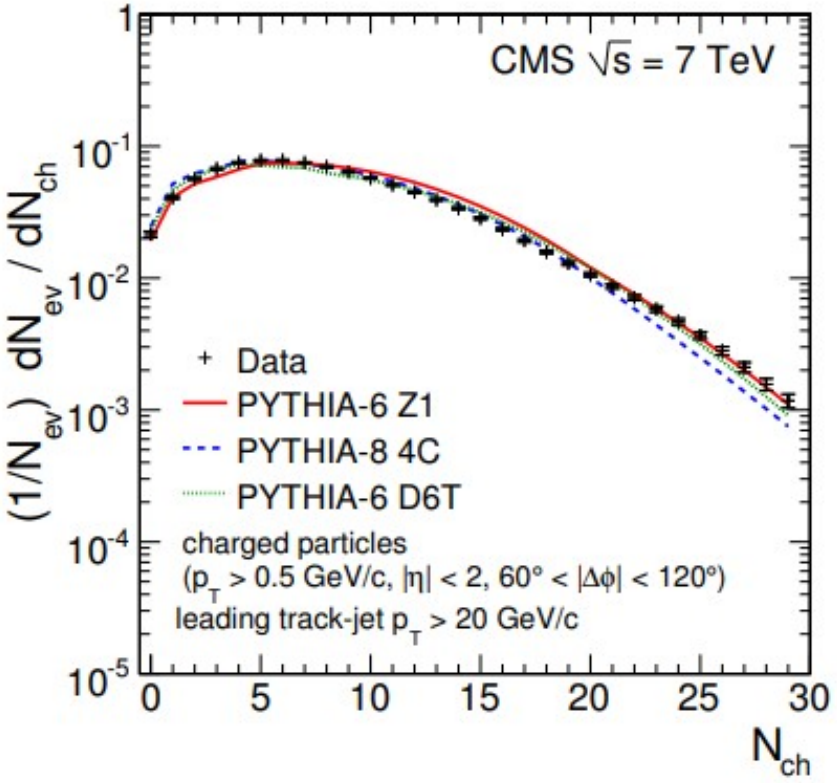
CPU resources accessed by the Test4Theory project (monthly averaged)



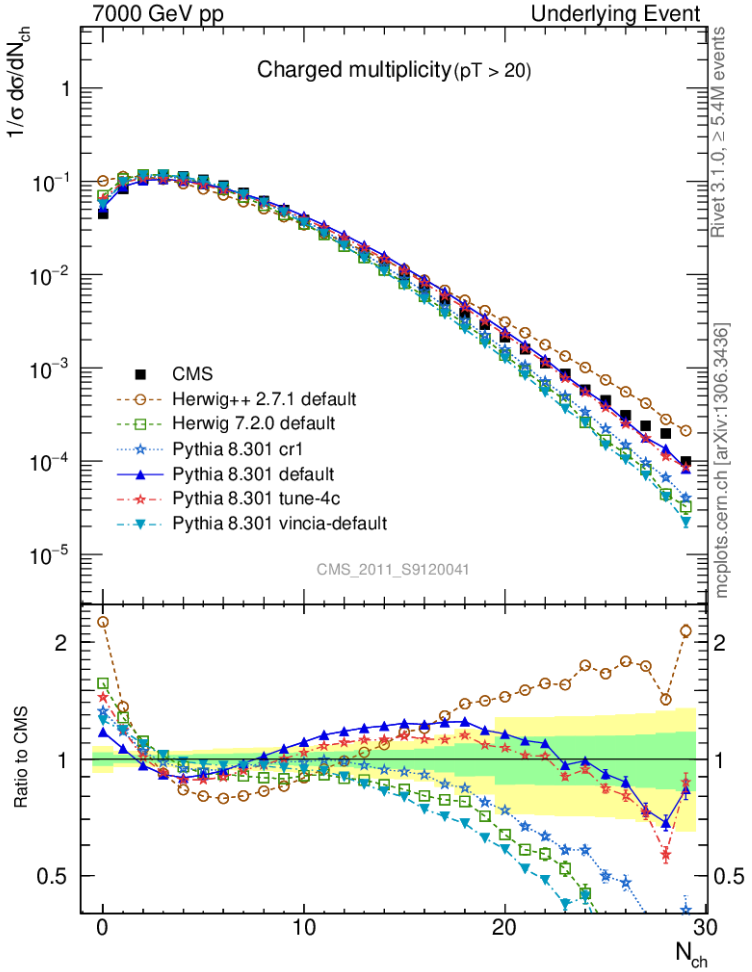
Plots

Generated distributions are collected on the MCPLOTS web server : dozens of generator-version-tune combinations for each data distribution ; they are plotted on the fly by a user's request.

Paper version



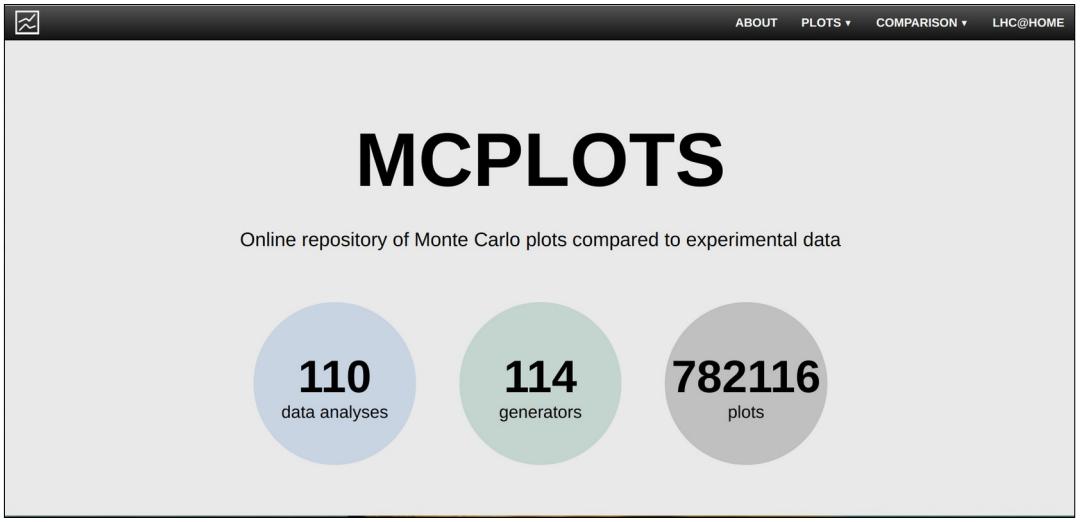
MCPLOTS version

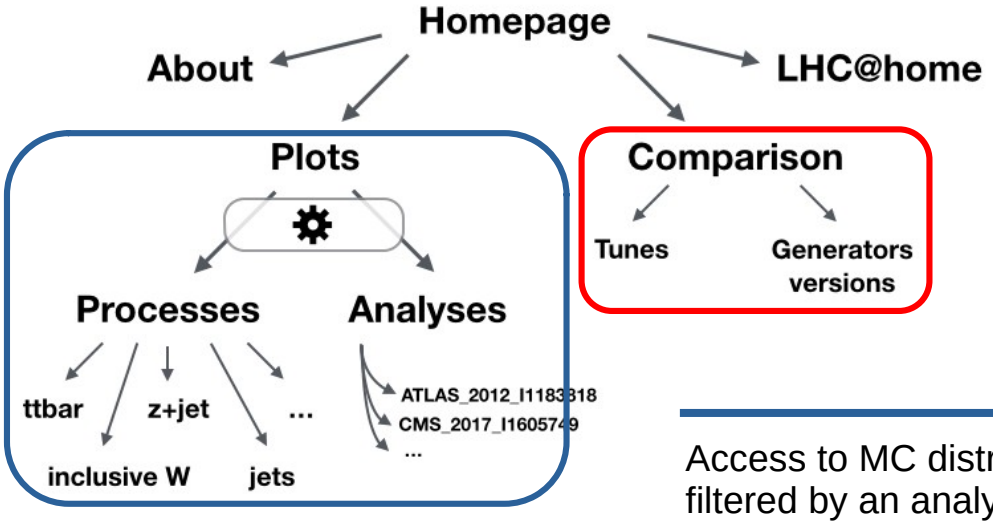


...or whichever generator-version-tune you want

Website

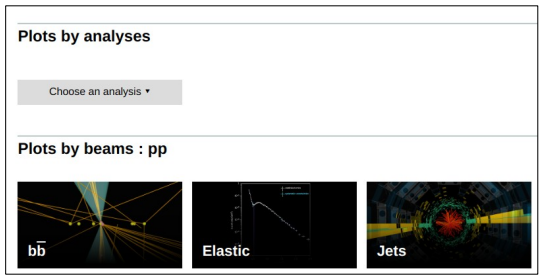
http://mcplots.cern.ch/





Numerical comparison of different **tunes** or **versions** of an MC generator

Access to MC distributions filtered by an analysis or a hard process

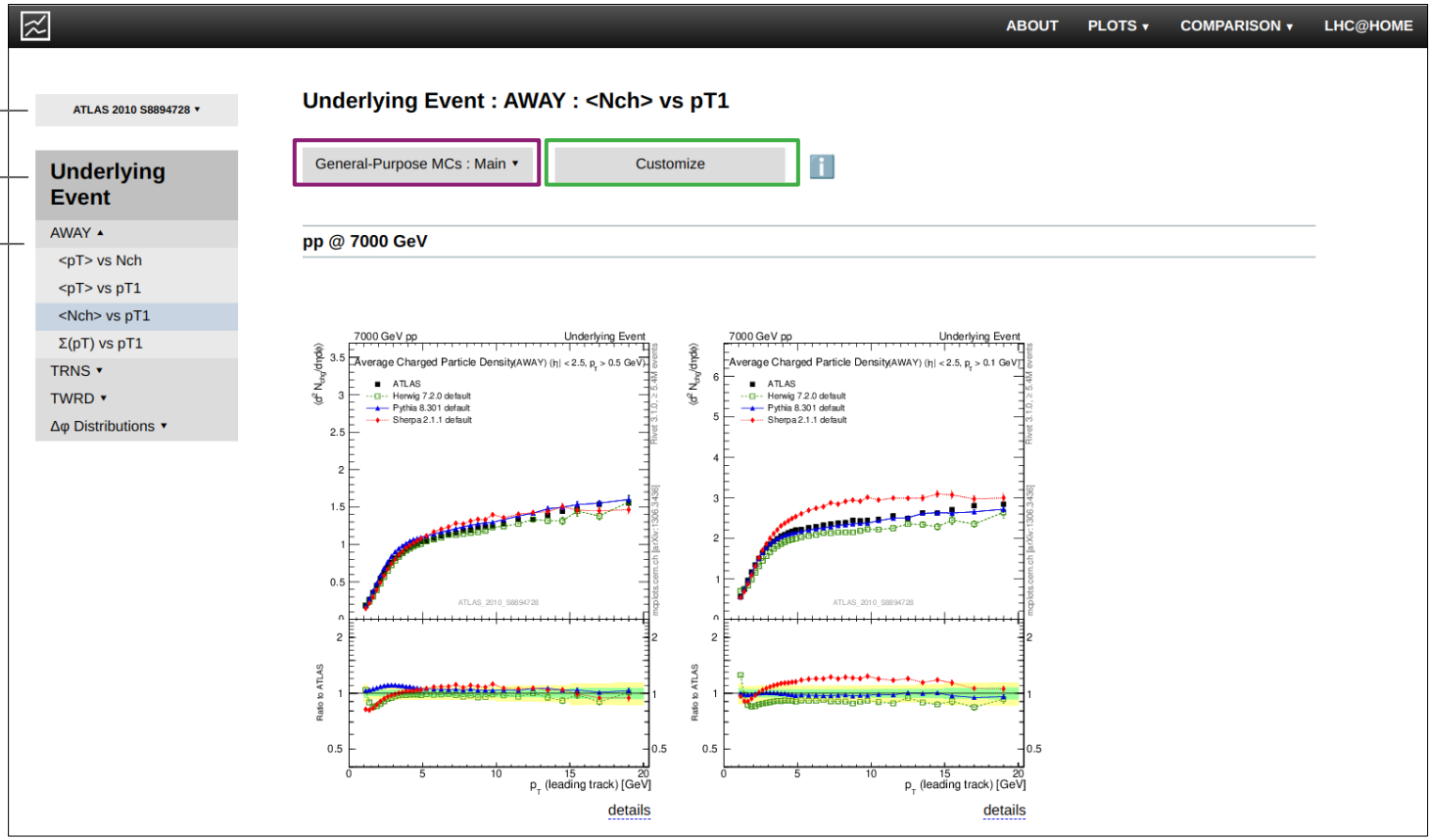


Website : plots

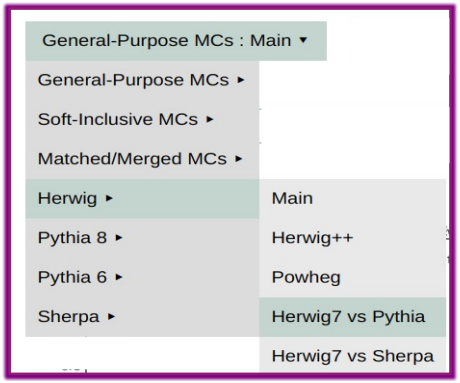
Analysis filter (if used)

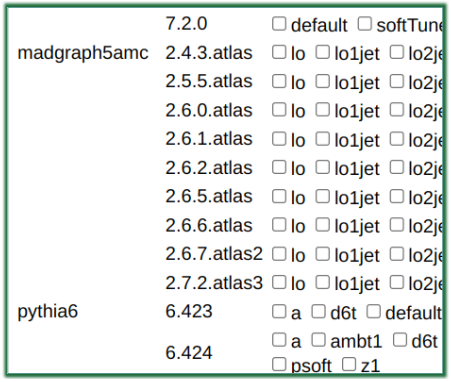
Hard process

Individual distributions



Possibility to choose what to plot either from **a pre-defined preset** or from **all MCs**





Website : plots

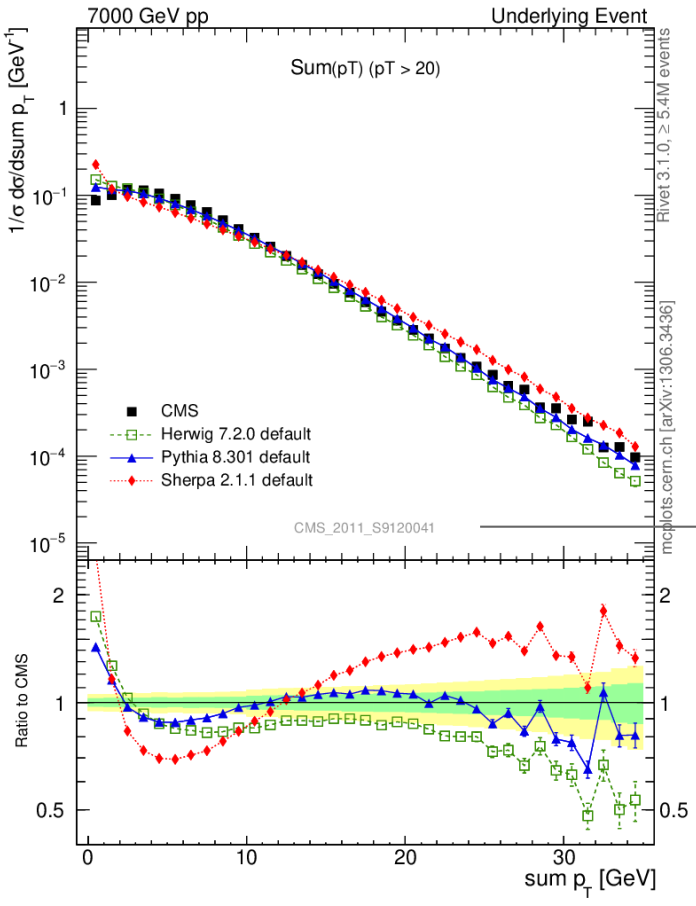
Beam parameters

Hard process

RIVET version and
# of MC events

Generator-version-tune
for each MC curve

RIVET reference



details

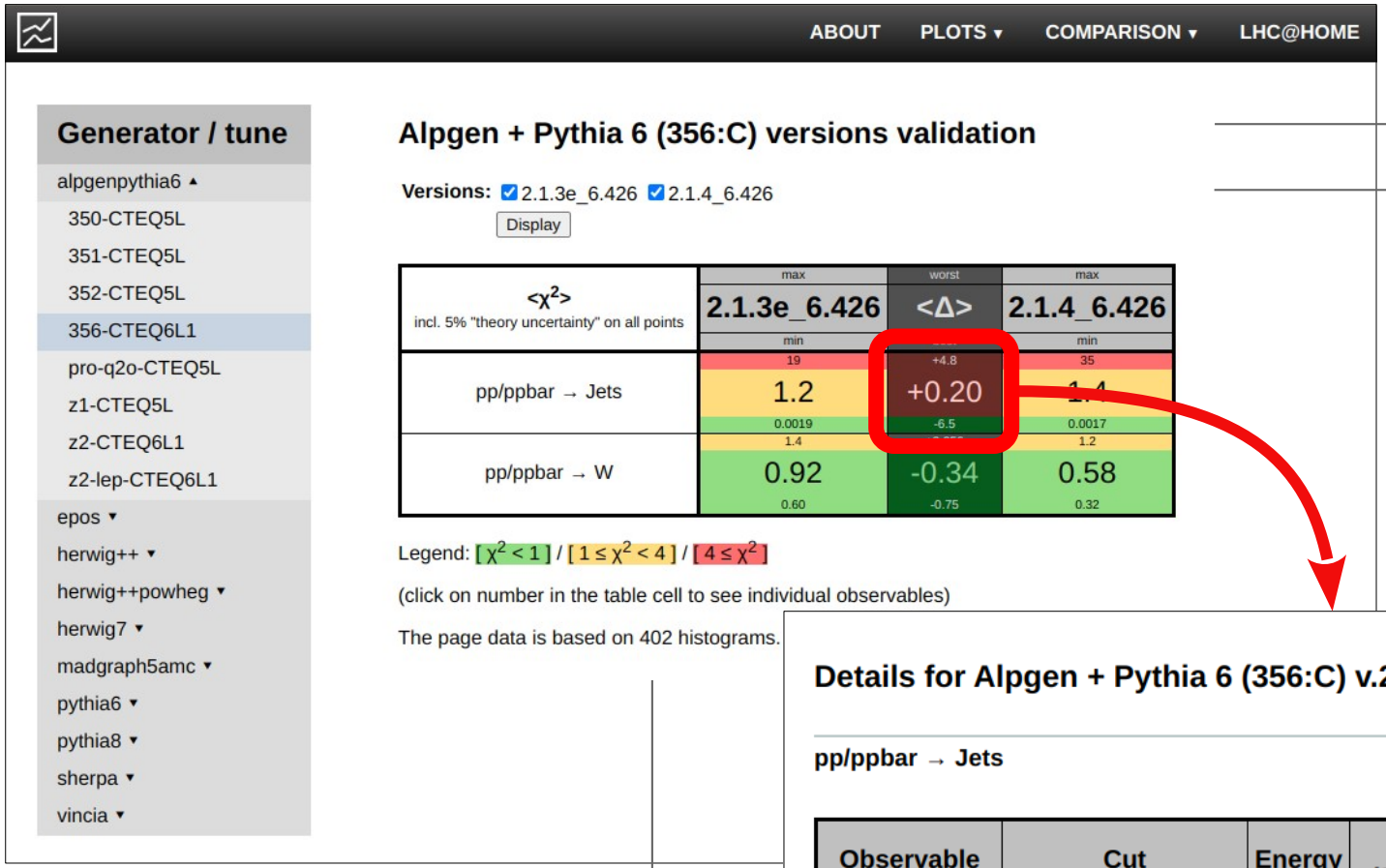
Download as: .pdf .eps .png .script.tgz #
CMS experiment: data | article paper
Herwig 7 (Def): data |generator card
Pythia 8 (Def): data |generator card
Sherpa (Def): data |generator card

Plot in higher resolution

Data distribution and article paper

Steering files and results
for each MC curve

Website : comparison



Generator (tune) name

Available versions to compare

χ² for individual observables

Number of distribution used to calculate χ²

Details for Alpgen + Pythia 6 (356:C) v.2.1.3e_6.426 vs. v.2.1.4_6.426

pp/ppbar → Jets

Observable	Cut	Energy	$\chi^2_{+5\%}$ (2.1.3e_6.426)	Δ	$\chi^2_{+5\%}$ (2.1.4_6.426)
23-jet Correlation	CMS 2013 (Forward)	7000	5.2	-3.8	1.4
	CMS 2013 (Central)	7000	5.1	-3.8	1.3
ET(J1)	CDF 1994	1800	n/a*	-	n/a*
ET(J2)	CDF 1994	1800	n/a*	-	n/a*
Transverse Minor	CMS 90 < pT < 125	7000	0.41	+0.49	0.90
	CMS 125 < pT < 200	7000	0.99	-	n/a*
	CMS pT > 200	7000	n/a*	-	n/a*
Transverse Thrust	CMS 90 < pT < 125	7000	0.72	+0.88	1.6
	CMS 125 < pT < 200	7000	0.72	-	n/a*
	CMS pT > 200	7000	1.1	-	n/a*

Current status

Always shown on the main page :

Implemented generators : Alpgen, Epos, Herwig++ and Herwig7, MadGraph, Pythia6 and Pythia8, Sherpa, Vincia

772 generator-version-tune combinations

110 data analyses with **1146 data distributions** implemented so far refer mostly to the ee and pp HEP collider experiments: ATLAS, CMS, D0 etc.

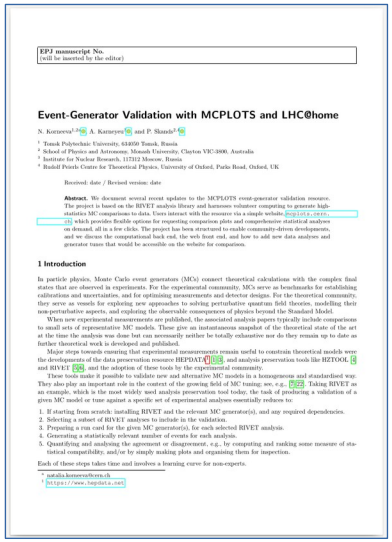
The repository is continously filled and the source code of the project is available :

https://gitlab.cern.ch/MCPLOTS/mcplots

Paper

The first paper was published ~10 year ago :
MCPLOTS: a particle physics resource based on volunteer computing
https://doi.org/10.1140/epjc/s10052-014-2714-9

The second one : **this monday**
Event-Generator Validation with MCPLOTS and LHC@home
https://arxiv.org/abs/2401.10621


EPJ manuscript No.
(will be inserted by the editor)

Event-Generator Validation with MCPLOTS and LHC@home

N. Korneeva[1,2] A. Karneyeu[3] and P. Skands[4]

[1] Tomsk Polytechnic University, 634050 Tomsk, Russia
[2] School of Physics and Astronomy, Monash University, Clayton VIC-3800, Australia
[3] Institute for Nuclear Research, 117312 Moscow, Russia
[4] Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Parks Road, Oxford, UK

Received: date / Revised version: date

Abstract. We document several recent updates to the MCPLOTS event-generator validation resource. The project is based on the RIVET analysis library and harnesses volunteer computing to generate high-statistics MC comparisons to data. Users interact with the resource via a simple website (mcplots.cern.ch) which provides flexible options for requesting comparison plots and comprehensive statistical analyses on demand, all in a few clicks. The project has been structured to enable community-driven developments, and we discuss the computational back end, the web front end, and how to add new data analyses and generator tunes that would be accessible on the website for comparison.

1 Introduction

In particle physics, Monte Carlo event generators (MCs) connect theoretical calculations with the complex final states that are observed in experiments. For the experimental community, MCs serve as benchmarks for establishing calibrations and uncertainties, and for optimizing measurements and detector designs. For the theoretical community, they serve as vessels for exploring new approaches to solving perturbative quantum field theories, modelling their non-perturbative aspects, and exploring the observable consequences of physics beyond the Standard Model.

When new experimental measurements are published, the associated analysis papers typically include comparisons to small sets of representative MC models. These give an instantaneous snapshot of the theoretical state of the art at the time the analysis was done but can necessarily neither be totally exhaustive nor do they remain up to date as further theoretical work is developed and published.

Major steps towards ensuring that experimental measurements remain useful to constrain theoretical models were the developments of the data preservation resource HEPDATA[1,2,3], and analysis preservation tools like RIVET tool [4] and RIVET [5,6], and the adoption of these tools by the experimental community.

These tools make it possible to validate new and alternative MC models in a homogeneous and standardised way. They also play an important role in the context of the growing field of MC tuning; see, e.g., [7,8,9]. Taking RIVET as an example, which is the most widely used analysis preservation tool today, the task of producing a validation of a given MC model or tune against a specific set of experimental analyses essentially reduces to:

1. If starting from scratch: installing RIVET and the relevant MC generator(s), and any required dependencies.
2. Selecting a subset of RIVET analyses to include in the validation.
3. Preparing a run card for the given MC generator(s), for each selected RIVET analysis.
4. Generating a statistically relevant number of events for each analysis.
5. Quantifying and analysing the agreement or disagreement, e.g., by computing and ranking some measure of statistical compatibility, and/or by simply making plots and organising them for inspection.

Each of these steps takes time and involves a learning curve for non-experts.

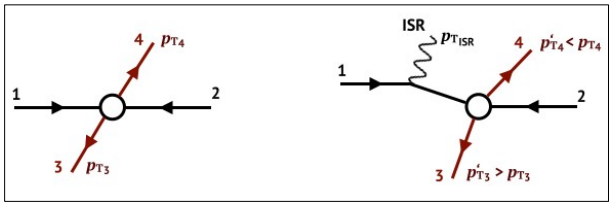
[*] natalia.korneeva@cern.ch
[**] https://www.hepdata.net

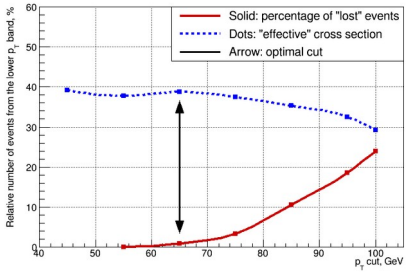
- Full description of the updated repository and database structure
- Comprehensive user's guide (the website functionality)
- Developer's guide : how to implement
  - ➔ a new data analysis
  - ➔ a new generator (version)
  - ➔ a new generator tune
- Phase-space cuts discussion

  Data analyses with high p_T jets in the final state : **physical** final state

  Generation cut on p_T of final state partons : hard **partonic** process

  Two methods to find an optimal generation cut are described in the new paper





Summary

- MCPLOTS : Online repository of Monte Carlo plots compared to experimental data
- CPU power: LHC@home
- Website: http://mcplots.cern.ch/
- Source code available to download

  https://gitlab.cern.ch/MCPLOTS/mcplots
- MCPLOTS paper

  https://arxiv.org/abs/2401.10621