





# A RIVETing journey: Analysis Preservation and Generator Tuning in High Energy Physics

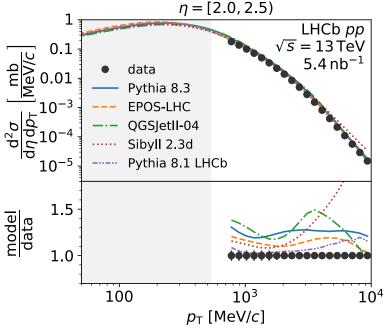
Johannes Albrecht, Hans Dembinski, Lars Kolk

Workshop on the tuning of hadronic interaction models

Wuppertal, 24.01.2024

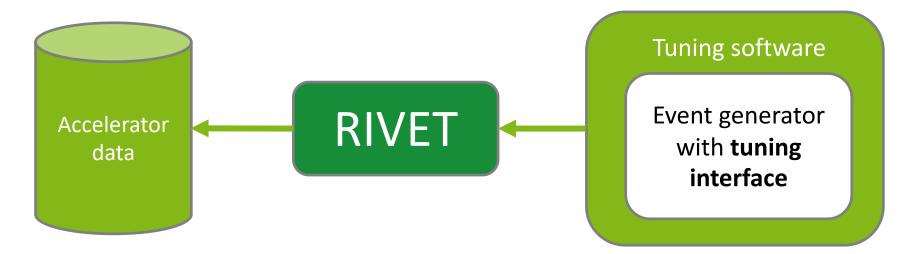
# **Motivation**

- Goal: Make paper results more accessible/usable
  - Make results publicly accessible
  - Provide results in machine readable format
  - Provide interface for generator tuning
- Archiving published analyses for public access
  - Promotes transparency and reproducibility
  - Increases impact of the analysis/paper
  - Improves utilisation of analysis for generator tuning
- Analysis preservation
  - 1. Data: HEPdata
  - 2. Analysis Code: <u>Rivet</u>



LHCb Collaboration, JHEP 01 (2022) 166

# HEP Generator tuning in a nutshell



Modified from H. Dembinski

## **Analysis Preservation Roadmap**





# HEPDATA

#### **HEPData**

- HEPData: Open-access repository used to preserve and share experimental data
  - Maintained by CERN
  - Huge catalogue of preserved analyses
  - Ensures long-term preservation of experimental results

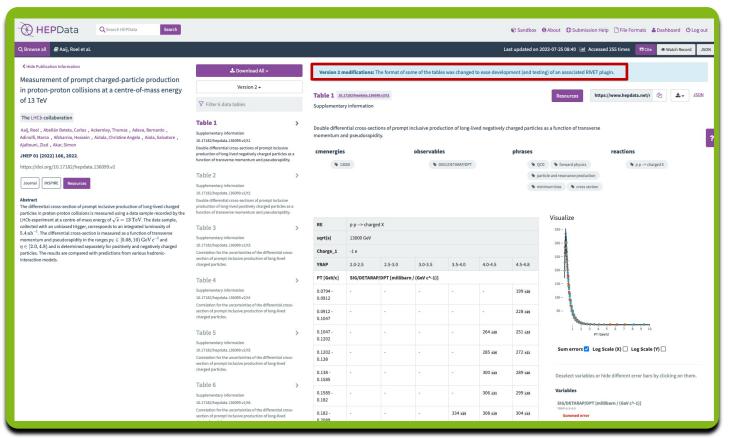
- HEPData Entries
  - Can be searched via many categories
  - Come with abstract and proper citation
  - Store data in tables

HEPData	O About  O Submission Help □ File Formats = 49 Sign in			
	Q Search HEPData Search Advanced 255N			
1 Max results - IF Sort by -	I Reverse order Showing 10 of \$2 results			
Date	• • <b>3 2 3</b>			
	[Ywwww] Neasurement of prompt charged-particle production in proton-proton collisions at a centre-of-mass energy of 13 TeV			
2010 2023	The LHCb collaboration Aaij, Roel ; Abellán Beteta, Carlos ; Ackernley, Thomas ; et al.			
Collaboration Reset	JHEP 01 (2022) 166, 2022.			
LHCB 82	ⓓ Inspire Record 1889335 % DOI 10.17182/hepdata.136099			
	The differential cross-section of prompt inclusive production of long-lived charged particles in proton-proton collisions is measured using a data sample recorded by the LHCb experiment at a centre-of-mass energy of $\sqrt{s} = 13  {\rm TeV}$ . The data sample, collected with an unbiased trigger, corresponds to an integrated luminosity of 5.4 mb <sup>-1</sup> . The differentia			
Subject_areas	experiment of a center of mass energy of $v_0 = 10$ TeV. The data sample, conected which an unbiased ungger, conesponds to an integrated unminosity of 0.4 ±0 . The dimeteriols			
hep-ex 82	Table 1 Double differential cross-sections of prompt inclusive production of long-lived negatively charged particles as a function of transverse momentum and			
nucl-ex 3	pseudorapidity.			
Phenomenology-HEP 1	Table 2 Double differential cross-sections of prompt inclusive production of long-lived positively charged particles as a function of transverse momentum and pseudorapidity.			
Phrases	Table 3 Correlation for the uncertainties of the differential cross-section of prompt inclusive production of long-lived charged particles.			
Proton-Proton Scattering 16	Yable 3 Contension for the uncertainties of the onlevendar cross-section of prompt inclusive production of long-wed charged particles.			
Inclusive 15	more			
Cross Section 10				
CM Energies (GeV)	[P Version 2] Measurement of antiproton production in $ m pHe$ collisions at $\sqrt{s_{NN}}=110$ GeV			
0.0 ≤√s < 1.0	The LHCb collaboration Aaij, Roel ; Abellán Beteta, Carlos ; Adeva, Bernardo ; et al.			
$1.0 \le \sqrt{s} < 2.0$ Phys.Rev.Lett. 121 (2018) 222001, 2018.				
2.0 ≤ √s < 5.0	2.0 ≤ √s < 5.0 📄 Inspire Record 1688924 % DOI 10.17182/hepdata.84584			
5.0 $\leq$ $\sqrt{s}$ < 10.0 The cross-section for prompt antiproton production in collisions of protons with an energy of 6.5 TeV incident on helium nuclei at rest is measured with the LHCb experiment from a				
10.0 ≤ √s < 100.0	data set corresponding to an integrated luminosity of 0.5 nb <sup>-1</sup> . The target is provided by injecting helium gas into the LHC beam line at the LHCb interaction point. The reported			
Next 5 Show All	⊞ 2 data tables			
Authors	Table 1 Double-differential cross-section in antiproton momentum and transverse momentum for antiproton production in collisions of 6.5 TeV protons on He nuclei at rest. The antiproton momentum is defined in the laboratory frame. Results are averaged over the given kinematic range of each bin. The uncertainty is split int			
Q Search authors	Table 1 Double-differential cross-section in antiproton momentum and transverse momentum for antiproton production in collisions of 6.5 TeV protons on He nuclei at			
Charpentier Philippe 53	rest. The antiproton momentum is defined in the laboratory frame. Results are averaged over the given kinematic range of each bin. The uncertainty is split int			

## **Example HEPData entry**

- HEPData Q. Search HEPData Search						🗑 Sandbox	🔁 About 🛛 🛱 Submissio	on Help 📑 File Formats 🌲 Dashboard 🖞 Log out
Q Browse all 🖉 Aaij, Roel et al.						Last updated on 2	023-05-26 13:02 네 Ac	cessed 254 times 🤧 Cite 👁 Watch Record JSON
C Hide Publication Information Measurement of prompt charged-particle production in proton-proton collisions at a centre-of-mass energy of 13 TeV The LHCb collaboration	Lownload All - Version 1 -	Sup	Table 1       H11ps://www.hepdata.net/r       Image: Complementary information         Differential cross-sections of prompt inclusive production of long-lived positively and negatively charged particles as a function of transverse momentum and pseudorapidity.       Image: Complementary information					
Anji, Roel , Abellán Beteta, Carlos , Ackernley, Thomas , Adeva, Bernardo , Adinolfi, Marco , Afsharnia, Hossein , Aldala, Chrístine Angela , Alola, Salvatore , Ajattouni, Ziad , Akar, Simon JHEP 01 (2022) 166, 2022. https://doi.org/10.17182/heppdata.136099.v1	Table 1       >         Resentery information       Simplify information (Research and Research and	cm	nenergies		observables	🗣 pa	D forward tticle and resonance nimum	reactions ◆ pp→charged X
wormal MSPRE Resources Abstract The differential cross-section of prompt inclusive production of long-lived charged particles in proton-proton collisions is measured using a data sample recorded by the LHC beoperiment at a centro-drmass energy of $\sqrt{c} = 13$ TeV. The data sample, collected with an unbased trigger, corresponds an integrated luminosity of $\delta$ . Ala $h^{-1}$ . The differential cross-section is measured as a function of transverse momentum and pseudospidity in the ranges $p_{T} \in [0.05, 0.05]$ (O eV $c^{-1}$ and $\eta \in [2.0, 4.8)$ and is determined separately for positively and negatively charged particles. The results are compared with predictions from various hadronic- interaction models.		Picket         Pp->charged X           sqrt(s)         pp->charged X           sqrt(s)         13000 GeV           charge_1         - 2.0           20-2.5         0.724 - 0.8318         8.978e1 (action of the sector)	Visualize	30- 9- 8- 8- 8- 8- 8- 8- 8- 8-				
		2.0 2.0 2.0	0 - 2.5 0 - 2.5 0 - 2.5	0.8318 - 0.955 0.955 - 1.0965 1.0965 - 1.2589 1.2589 - 1.4454	6.846e+1 x3.78400 5.058e+1 x2.77400 3.619e+1 x3.78400 2.515e+1 x3.78440		3- 2- 2022242630033445364052446 EAMP 0.00017 2043000 Brushing Enabled? □ X Axis ETARAP ∨ Y Axis PT (GeW/C) ∨	ETARAP
		2.0	0 - 2.5	1.4454 - 1.6596 1.6596 - 1.9055 1.9055 - 2.1878 2.1878 - 2.5119	1.6941e+1 ±324e-1 1.1036e+1 ±516e-1 6.957e+0 ±2.44e-1 4.232e+0 ±138-1			ar
		2.0	0 - 2.5	2.5119 - 2.884 2.884 - 3.3113 3.3113 - 3.8019	2.4854e+0 16.31e-2 1.4032e+0 16.31e-2 7.631e-1 16.88e-2			

### **Example HEPData entry**



Lars Kolk

# **HEPData-Submissions**

- Each submission consists of ...
- One submission.yaml
  - Links to every table in the submission
  - Provides metadata for the HEPData entry
- One .yaml file for every table in the entry
  - Contain data points, errors
- Refer to <u>documentation</u> or previously uploaded entries : )
- Contact your <u>coordinator</u> to initiate a submission
- Use <u>hepdata-validator</u> or <u>HEPData sandbox</u> to validate your files

	1 submission.yaml — converted_data
EXPLORER	(.) submission.yaml ×
CONVERTED_DATA L L C A CONVERTED_DATA L C C A CONVERTED_DATA L C C C C C C C C C C C C C C C C C C	<pre>1</pre>
	29 data_file: cross_section_0.yaml 30
	, 1, Col 1 Spaces: 2 UTF-8 LF {} YAML

# **HEPData-Submissions**

- Each submission consists of ...
- One submission.yaml
  - Links to every table in the submission
  - Provides metadata for the HEPData entry
- One .yaml file for every table in the entry
  - Contain data points, errors
- Refer to <u>documentation</u> or previously uploaded entries : )
- Contact your <u>coordinator</u> to initiate a submission
- Use <u>hepdata-validator</u> or <u>HEPData sandbox</u> to validate your files

¢	EXPLORER ····	, cross_section_0.yaml ×	
5	~ CONVERTED_DATA 自日のひ日	LHCB_2021_I1889335 > () cross_section_0.yaml > [ ] dependent_variables	
	✓ ☐ LHCB_2021_I1889335	1 dependent_variables:	
	Cross_section_0.yaml	2 - header: {name: SIG/DETARAP/DPT, units: millibarn / (GeV	
	() cross_section_1.yaml		
	. cross_section_correlation_0.yaml	4 - {name: RE, value: p p> charged X}	
	<pre>{} cross_section_correlation_1.yaml</pre>	5 - {name: sqrt(s), units: GeV, value: 13000}	
	cross_section_correlation_2.yaml	6 - {name: Charge_1, units: e, value: -1}	
	<pre>{} cross_section_correlation_3.yaml</pre>	7 - {name: YRAP, value: 2.0-2.5}	
7 3		8 values:	
	LHCB_2021_I1889335.tar.gz	9 - {value: '-'} 10 - {value: '-'}	
5	CHOB_2021_110093333.tal.gz	10 - {value: - }	
		12 - {value: - }	
		13 - {value: '-'}	
5		14 - {value: '-'}	
		15 - {value: '-'}	
		22 - {value: '-'} 23 - {value: '-'}	
		23 - {value:}	
		25 - value: '90'	
		26 errors:	
		27 - {symerror: '5'}	
		28 - value: '68'	
		30 - {symerror: '4'}	
		31 - value: '50.6'	
		33 - {symerror: '2.8'}	
	> OUTLINE	34 - value: '36.2'	
2		35 errors: 36 - {symerror: '2.0'}	
	> TIMELINE	36 - {symerror: '2.0'}	

# RIVET

#### **Rivet**

- <u>Rivet</u>: Most widespread way by which analysis code from the LHC and other high-energy collider experiments is preserved.
  - Used for generator tuning
  - Provides set of experimental analyses
  - Rivet <u>wishlist</u>, <u>installation</u>
- Analysis name: <Experiment>\_<publication\_year>\_I<InspireID>
  - E.g.: LHCB\_2021\_I1889335
- Each Rivet-analysis consists of one ...
  - .cc file containing the analysis code
  - .yoda file containing the data points (from HEPData)
  - .info file containing metadata
  - .plot file setting plotting options

📥 Download All 🗸	
YAML with resource files	
YAML	
YODA	
ROOT	
CSV	>

- An analysis is usually written in one .cc file containing
  - A no-argument constructor
  - Three analysis event loop methods:
    - 1. init: book histograms, initialise counters, etc.
    - 2. analyze: per event: select particles, apply cuts, construct observables, fill histograms etc.
    - 3. finalize: normalize/scale/divide histograms, tuples, etc.
  - A minimal hook into the plugin system
- Projections: observable calculator objects that are called by the analysis' apply(event) method.
  - Refer to projection declared in init method
- Compiled via rivet-build <NAME>.cc
  - Generates a RivetAnalysis.so file



•	C+ LHCB_2021_I1889335.cc ₃ ×					
	C++ LHCB_2021_I1889335.cc >					
	1 // -*-C++ - *-					
	2 #include "Rivet/Analysis.hh"					
	<pre>3 #include "Rivet/Projections/AliceCommon.hh" 4</pre>					
	5 namespace Rivet					
	6 {					
	7 8 /// @brief Inelastic section in pp collisions at 13 TeV for charged particles in LHCb acceptance 9 class LHCB_2021_I1889335 : public Analysis					
	10 {					
•	11 public:					
	12 /// Constructor					
	<pre>13 RIVET_DEFAULT_ANALYSIS_CTOR(LHCB_2021_I1889335);</pre>					
	14					
	15 /// @name Analysis methods					
•	16 //@{					
	17					
	18 /// Book histograms and initialise projections before the run					

	10	777 book histograms and initiate projections before the ran
•	19	void init()
	20	{
	21	
	22	<pre>// Register projection for primary particles</pre>
	23	<pre>declare(ALICE::PrimaryParticles(Cuts::etaIn(ETAMIN, ETAMAX) &amp;&amp; Cuts::abscharge &gt; 0), "APRIM");</pre>
	24	
	25	{HistolDPtr tmp; _h_ppInel_neg.add(2.0, 2.5, book(tmp, 1, 1, 1));}
	26	{HistolDPtr tmp; _h_ppInel_neg.add(2.5, 3.0, book(tmp, 1, 1, 2));}
	27	{HistolDPtr tmp; _h_ppInel_neg.add(3.0, 3.5, book(tmp, 1, 1, 3));}
	28	{HistolDPtr tmp; h ppInel neg.add(3.5, 4.0, book(tmp, 1, 1, 4));}
	29	{HistolDPtr tmp; _h_ppInel_neg.add(4.0, 4.5, book(tmp, 1, 1, 5));}
	30	{HistolDPtr tmp; _h_ppInel_neg.add(4.5, 4.8, book(tmp, 1, 1, 6));}
•	31	
	32	<pre>{HistolDPtr tmp; _h_ppInel_pos.add(2.0, 2.5, book(tmp, 2, 1, 1));}</pre>
	33	<pre>{HistolDPtr tmp; _h_ppInel_pos.add(2.5, 3.0, book(tmp, 2, 1, 2));}</pre>
	34	<pre>{HistolDPtr tmp; _h_ppInel_pos.add(3.0, 3.5, book(tmp, 2, 1, 3));}</pre>
	35	<pre>{Histo1DPtr tmp; _h_ppInel_pos.add(3.5, 4.0, book(tmp, 2, 1, 4));}</pre>
	36	<pre>{Histo1DPtr tmp; _h_ppInel_pos.add(4.0, 4.5, book(tmp, 2, 1, 5));}</pre>
٠	37	<pre>{Histo1DPtr tmp; _h_ppInel_pos.add(4.5, 4.8, book(tmp, 2, 1, 6));}</pre>
	38	n → primer de la constanti de constante de la constanta da constanti de la Cidada da Cidada da Cidada da Cidad

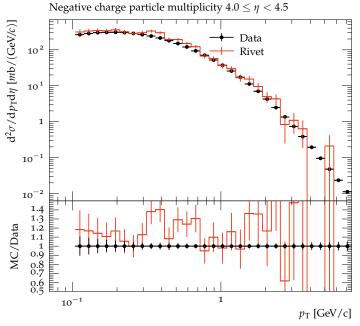
```
void analyze(const Event &event)
                 const Particles cfs = apply<ALICE::PrimaryParticles>(event, "APRIM").particles();
                 for (const Particle& myp : cfs)
                   if (myp.charge() < 0)
                     _h_ppInel_neg.fill(myp.pseudorapidity(), myp.momentum().pT());
.
                     _h_ppInel_pos.fill(myp.pseudorapidity(), myp.momentum().pT());
•
          Generates a riverAnatysis.so nie
```

```
void finalize()
    const double scale factor = crossSection() / millibarn / sumOfWeights();
    std::vector<double> binWidths = {0.5, 0.5, 0.5, 0.5, 0.5, 0.3};
    for (size t i = 0; i < binWidths.size(); i++)</pre>
      _h_ppInel_neg.histos()[i]->scaleW(scale_factor / binWidths[i]);
      h ppInel pos.histos()[i]->scaleW(scale factor / binWidths[i]);
  BinnedHistogram h ppInel_neg;
  BinnedHistogram h ppInel pos;
  const double ETAMIN = 2.0, ETAMAX = 4.8;
RIVET DECLARE PLUGIN(LHCB 2021 I1889335);
```

•

# Running and submitting a Rivet-plugin

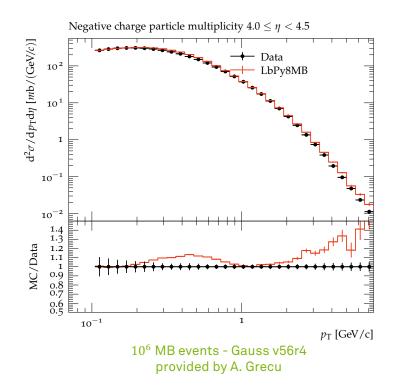
- You can run your compiled code via rivet -a <Name> <inputfile>
  - inputfile: .hepmc file
  - Style: .plot file
  - Rivet provides <u>example files</u> to test on
  - Alternative: Let your favourite generator generate some events for you!
  - Produces Rivet.yoda file
- Plotting: rivet-mkhtml --errs Rivet.yoda
- Outputs look sane → Internal Review /MR in <u>RIVET</u>



10<sup>3</sup> MB events - LHC-13-Minbias.hepmc.gz

# Running and submitting a Rivet-plugin

- You can run your compiled code via rivet -a <Name> <inputfile>
  - inputfile: .hepmc file
  - Style: .plot file
  - Rivet provides <u>example files</u> to test on
  - Alternative: Let your favourite generator generate some events for you!
  - Produces Rivet.yoda file
- Plotting: rivet-mkhtml --errs Rivet.yoda
- Outputs look sane → Internal Review /MR in <u>RIVET</u>



# **Additional Remarks**

# **Analysis Preservation Guide**

• Written <u>Analysis Preservation Guide</u>, including all important steps

LHCb Simulation Project		Work Packages Useful Information Edit on GitLab <sup>II</sup>				
Quick search Search	Docs » Analysis Preservation with HEPData and Rivet	← WP R - Machine Induced Background & Radiation   Information for Liaisons →				
Work Packages WP G - Generators WP D - Detector Modeling WP F - Framework & New Technologies WP P - Monte Carlo Productions WP S - Detailed and Fast Simulation WP R - Machine Induced Background & Radiation	The goal is to preserve analyses in such a way that an interface for the generator tuning community. T standardized and machine-readable format on HEF tuning community, which includes applied cuts and This document describes the process of preserving	Analysis Preservation with HEPData and Rivet The goal is to preserve analyses in such a way that the data is easily accessible and reproducible, while providing an interface for the generator tuning community. This can be achieved by storing the results of the analysis in a standardized and machine-readable format on HEPData while using Rivet to provide an interface for the generator tuning community, which includes applied cuts and selections of the analysis. This document describes the process of preserving an analysis by the example of the analysis LHCB-PAPER- 2021-010, based on a presentation given by Lars Kolk in the Simulation Meeting.				
Useful Information	HEPData					
Analysis Preservation with HEPData and Rivet HEPData Rivet Information for Liaisons	data-points and tables of a published paper in a machine community and is hosted by CERN. The website of	serve analyses in high energy physics. It is used to store chosen achine-readable way. The service provided by the HEP fers a huge catalogue of preserved analyses already and ensures ch analysis is stored in a record, which is a collection of files ts of the following files:				

Lars Kolk

# Connection to SFB1491

- F3 & F4 Deliverables: "Selected LHCb analyses will have been adopted for the use with Rivet"
- Provide plugins for generator tuning in (astro)-particle physics
- Several rivet plugins written by Chloé, e.g.
  - <u>NA61SHINE\_2019\_I1754136</u>
  - <u>HARP\_2008\_1778842</u>
  - <u>HARPCDP\_2010\_1863735</u>
  - <u>LHCB\_2011\_I891233</u>
- Pion/Kaon/Proton projectile against carbon target in forward rapidity regions

# Summary

- Archiving published analyses for public access
  - Promotes transparency and reproducibility
  - Increases impact of analyses
- Analyses can be archived ...
  - Data: HEPData
  - Code: Rivet
- Preservation procedure shown by example LHCB\_2021\_I1889335
- Written guide to help with analysis preservation
- Rivet plugins are important deliverables for the SFB
- Rivet plugins can be used by **anybody's** generators based off **your** results!