

CEDAR tools update

What's new in Rivet 3, HepMC 3 & LHAPDF 6

Andy Buckley, University of Glasgow
for the Rivet, HepMC3 and LHAPDF teams

21st MCnet meeting
10 September 2020



University
of Glasgow



THE ROYAL
SOCIETY



Rivet 3

❖ Rivet v3 released June 2019: currently 3.1.2, July 2020

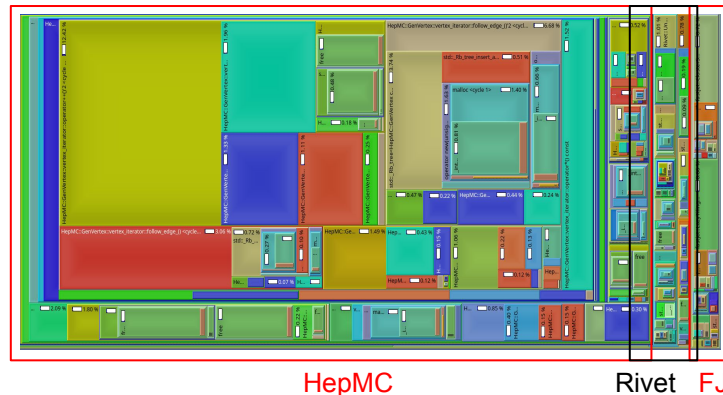
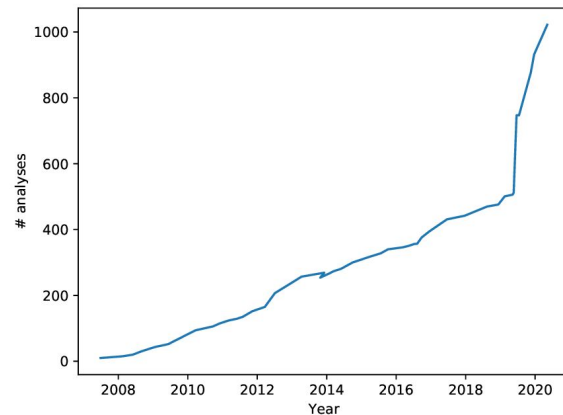
- automatic MC systematics multiweight handling
- heavy ion machinery, analysis parameters, ...
- explosion in e^+e^- hadron analyses: thanks, Peter R!
- paper: <https://arxiv.org/abs/1912.05451>
- **~1000 analyses now bundled!**
- Docker images for rivet and rivet+\$generator

❖ Plus: YODA stats library and HepData output augmented to include bin correlation data

(Louie Corpe, Graeme Watt, AB)

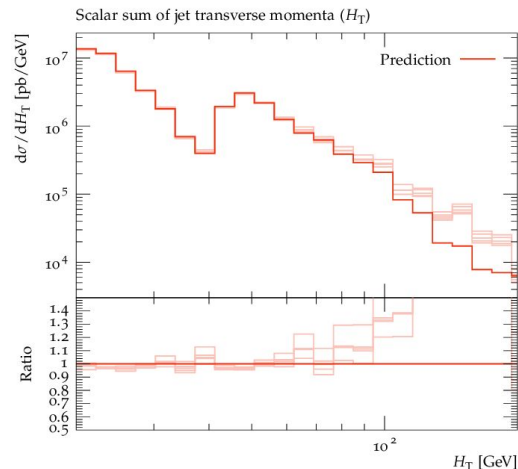
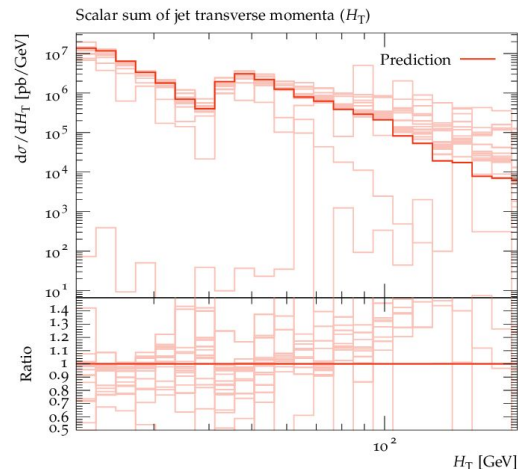
❖ Recent activities: migration to gitlab.com for analysis integration (MRs) and CI.

Docs improvements. Stats, plotting and performance focus groups + GSoC students



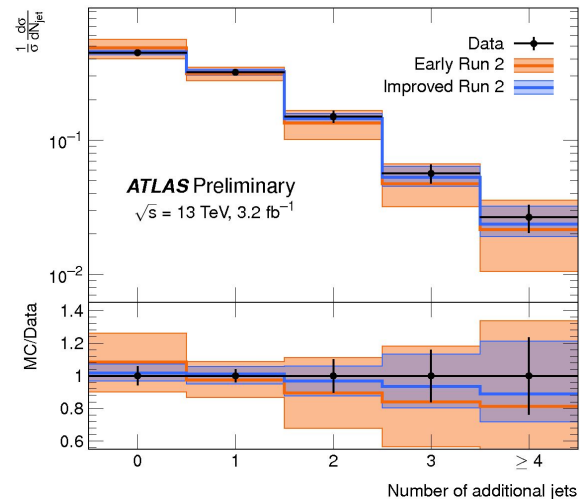
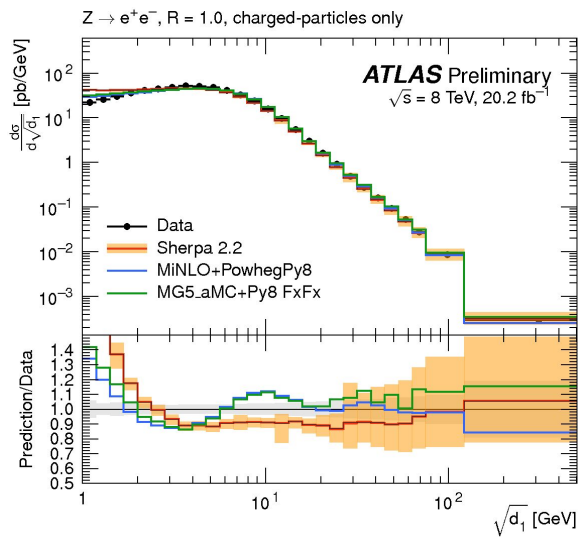
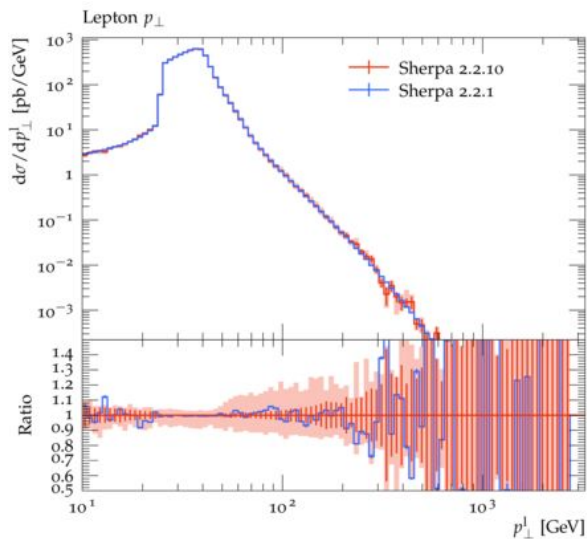
Rivet multiweights

- ❖ **Very (internally) complex automatic handling of MC systematics weight vectors**
- ❖ **Complex because ~invisible to users: data objects *look* like histograms etc. but are secretly multiplexed.** Also handles counter-events and many histo types: many levels of abstraction
- ❖ **Can re-run finalisation calculation with combined runs: RAW histogram stage.**
Best practice evolving, e.g. **weight ignoring & filtering**, efficiency improvements ongoing.
Weight-name standardisation: see discussion
- ❖ **HepMC 2.06.11 / R3.1.2 weight-name order bugfix!**



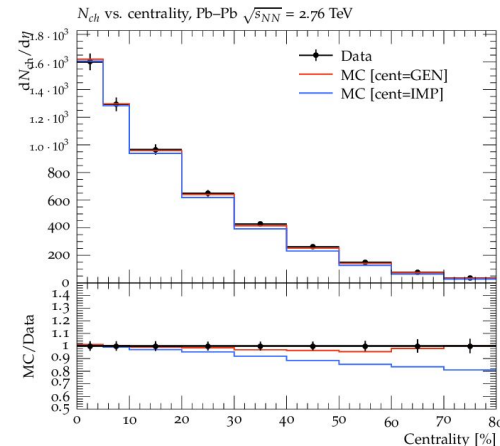
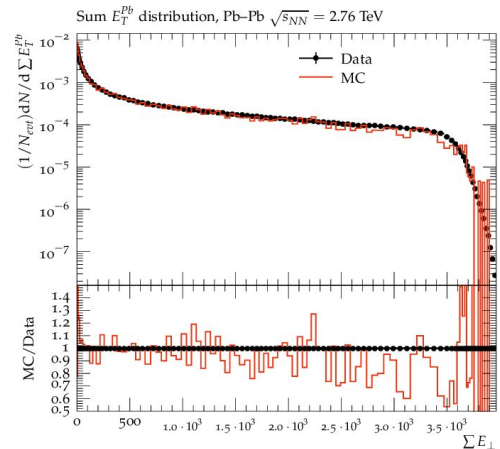
Rivet multiweights in action

ATLAS MC studies have been a significant driver of this feature (thanks to Chris Gutschow)



Rivet heavy ion

- ❖ “Adding heavy ion support” sounds trivial!
- ❖ Actually a stern test, with far-reaching impacts. HI observables often require centrality calibration curves: need a 2-pass run. And event/event correlations... centrality-binned!
- ❖ Also swappable definitions: few HI generators are general-purpose. Thrashed out through e.g. Aug 2017 NBI workshop (Christian Bierlich et al + ALICE reps)
- ❖ Paper: <https://arxiv.org/abs/2001.10737>
Also spurring feedback to HI MC standards



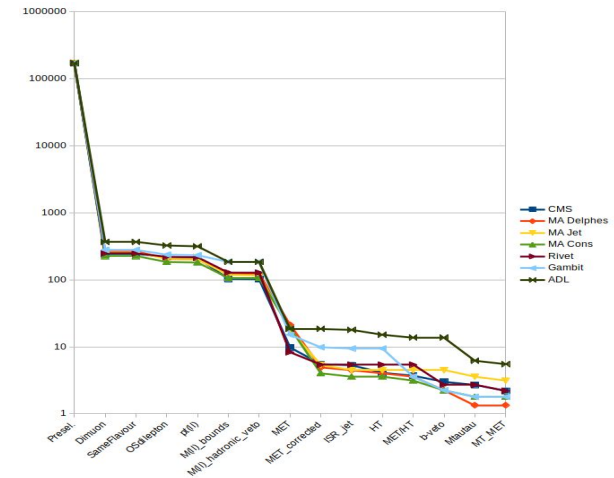
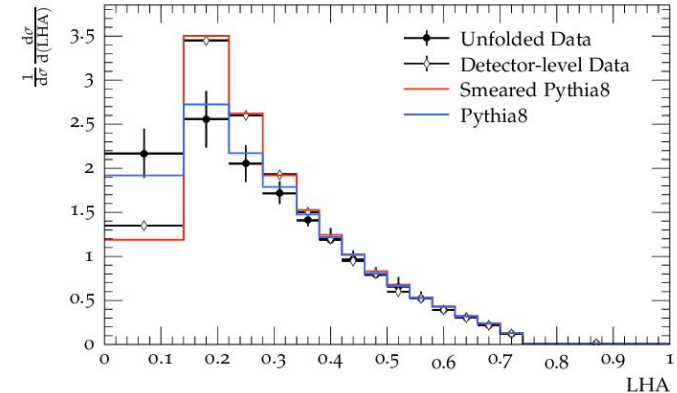
Rivet and BSM search recasting

❖ Detector smearing system:

- developed based on Gambit experience (AB)
- key features cf. Delphes, but more flexible & more analysis-specific
- Paper: <https://arxiv.org/abs/1910.01637>
(including “tuned” *jet-substructure smearing*)

❖ Same speed as Delphes via HepMC; approach to be repeated in new universal recast code

- ## ❖ Involved in Les Houches 2019 cutflow comparisons and global fit tests. Performance very good! Important cross-check on established recast codes



Rivet analysis coverage and compliance

- ❖ Experiment published-analysis coverage tracked at e.g. <https://rivet.hepforge.org/rivet-coverage#cmsexpt>
- ❖ CMS lags ATLAS on core SM analysis coverage: 38% vs 55% on “core” measurement analyses
Impact via e.g. Contur <https://arxiv.org/abs/2006.07172>
- ❖ Critical issue: analyses submitted with reference data incompatible with that submitted to HepData
- ❖ Rivet needs to be able to synchronise analysis references with latest data: tighter QA, technical evasive manoeuvres like sync-patching scripts

Rivet analysis coverage (no searches, no heavy ion)

Rivet analyses exist for 845/4241 papers = 20%. 153 priority analyses required.

Total number of Inspire papers scanned = 7280, at 2020-07-02

Breakdown by identified experiment (in development):

Key	ALICE	ATLAS	CMS	LHCb	Forward	HERA	$e^+e^- (\geq 12 \text{ GeV})$	$e^+e^- (\leq 12 \text{ GeV})$
Rivet wanted (total):	72	111	126	153	43	461	765	647
Rivet REALLY wanted:	17	42	61	9	0	13	1	3
Rivet provided:	14/86 = 16%	135/246 = 55%	77/203 = 38%	13/196 = 7%	8/51 = 16%	9/470 = 2%	166/931 = 18%	344/991 = 35%

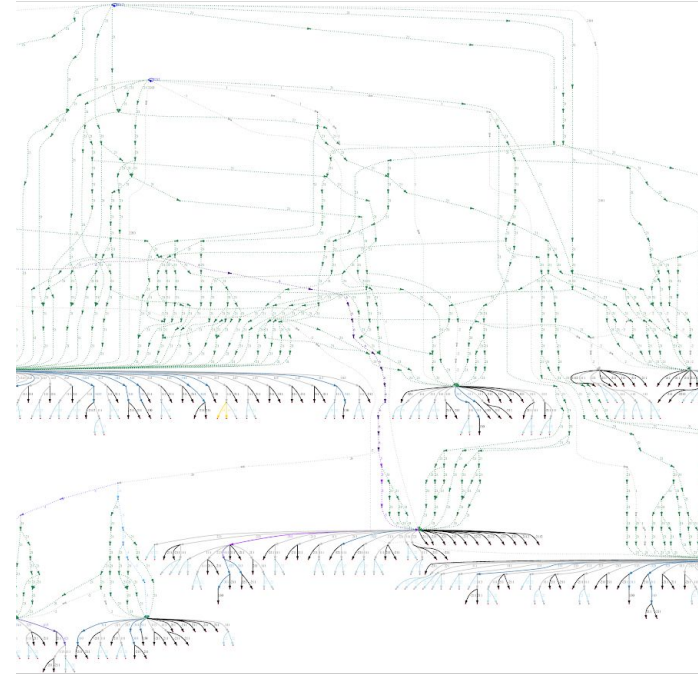
Show graylist Show blacklist

ALICE ATLAS CMS LHCb Forward HERA $e^+e^- (\geq 12 \text{ GeV})$ $e^+e^- (\leq 12 \text{ GeV})$ Tevatron RHIC SPS Other

ATLAS: Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS experiment Inspire ID: 1802524 arXiv ID: 2006.13076 Report ID: CERN-EP-2020-096 Links: Inspire arXiv
ATLAS: Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ using the ATLAS detector Inspire ID: 1801434 arXiv ID: 2006.09274 Report ID: CERN-EP-2020-063 Links: Inspire CDS arXiv
ATLAS: Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13 \text{ TeV}$ Inspire ID: 1790439 arXiv ID: 2004.03969 Report ID: CERN-EP-2020-035 Links: Inspire CDS arXiv HepData ATLAS_2020_11790439
ATLAS: Measurement of the Lund jet plane using charged particles in 13 TeV proton-proton collisions with the ATLAS detector Inspire ID: 1790256 arXiv ID: 2004.03540 Report ID: CERN-EP-2020-030 Links: Inspire DOI/journal CDS arXiv HepData ATLAS_2020_11790256
ATLAS: Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector Inspire ID: 1788444 arXiv ID: 2003.11960 Report ID: CERN-EP-2020-022 Links: Inspire CDS arXiv
ATLAS: Measurement of isolated-photon plus two-jet production in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector Inspire ID: 1772071 arXiv ID: 1912.09866 Report ID: CERN-EP-2019-210 Links: Inspire CDS arXiv
ATLAS: A measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} = 13 \text{ TeV}$

HepMC 3

- ❖ **HepMC2 was an unmaintainable mess!** Ownership, constness, limitations in physics content, defaults...
- ❖ **Now: arbitrary attributes, sane constness, LHE embedding.** See <https://arxiv.org/abs/1912.08005>
- ❖ Where development will happen: lots of computational bottlenecks have been in HepMC, **v3 is a consistent platform in which to address them, and hone the interface**
- ❖ E.g. Holger Schulz work on HDF5 output format
- ❖ **Could we agree that SHG generators output a “minimal” (non-debug) HepMC record by default?**



LHAPDF 6

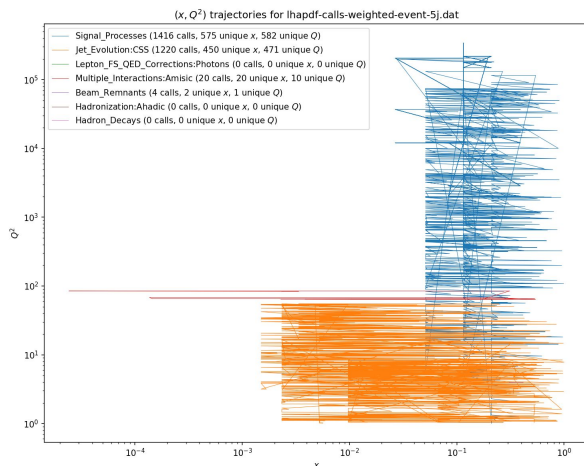
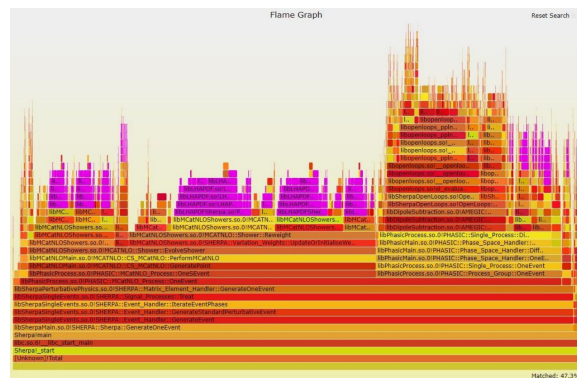
❖ LHAPDF performance studies via UK ECHP

❖ New 6.3.0 release with a thread-safe cache across compatible PDF grids (including all flavours)

Generator re-engineering needed to make best use of it: esp. across $O(100)$ systematic variations in a single PDF set

❖ Precision also an issue: trade-off between pluggability (e.g. a higher-order spline interpolator) and performance. LHAPDF 7 step back to a less modular system, now with single data format?

❖ Short-term studentship potential for all the above



Summary

- ❖ Lots of MC tools development under the CEDAR project: use the new Rivet, HepMC, LHAPDF, Professor, Contur
- ❖ Rivet 3 is now the standard: multiweights, heavy ion analysis, many more analyses, better MR...

Multiweights essential: analysis performed only once, weights invisibly propagated. **Need standards!**
Perfect run combination via rivet-merge

- ❖ HepMC 3: now full SHG support? Can we reconsider how generators write events: how much is essential, how much is debug?
- ❖ LHAPDF: PDF errors are a huge CPU sink for NLO events. Structuring calculations for efficiency



Professor and correlations

- ❖ Professor method for tuning well-established
- ❖ Holger Schulz (FNAL) recently released a high-performance update as “Apprentice”
<https://iamholger.gitbook.io/apprentice/>
- ❖ CEDAR project (AB, HS) on a better definition of MC tuning uncertainties: use old “eigentunes” idea, but define tolerances from statistical bootstrap rather than classic χ^2 threshold.
- ❖ Then perform uncertainty pruning/ranking based on a feature sensitivity measure.
- ❖ Papers soon

