

MC toolchain lessons from the LHC

From supporting tools to MC production

Andy Buckley,
University of Glasgow

MC4EIC
IPPP Durham, 4 June 2024



University
of Glasgow

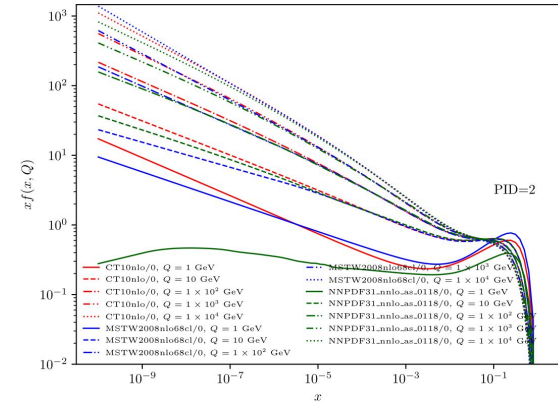
LHC experience and the MC toolchain

- ❖ The LHC has (among a few other things) seen huge developments in the sophistication of the MC ecosystem
 - As will be discussed, the physics content of the generators is a world away from pre-LHC:
 - NLO, multileg matching & merging, weighted variations, ...
 - But also the “supporting tools”:
 - Event formats: HepMC, LHE, HDF5LHE, ...
 - Parton densities: LHAPDF universal format and interpolators
 - Validation: HZTOOL → Rivet
 - Tuning: Professor and Apprentice toolkits plus e.g. Autotune
 - BSM interpretation, detector fast-sim, generative ML... *not for today!*
- ❖ I'll give a **whistle-stop tour** of what's available
- ❖ And also some **practical, logistical and social lessons learned**

Before the MC: LHAPDF parton density library

- ❖ PDFs are a crucial input to MC
 - Also a key *output* of ep/eA colliders, of course...
 - Need a functional library of (~1500...) established fits
 - Now standard to expect a *error set* for each fit

- ❖ LHAPDF
 - Library evolved out of Les Houches workshops ~ 2001
 - Backward compatibility with original CERN PDFLIB
 - Gradually extended for error sets, but crumbling F77 and patchwork of incompatible submitted evolution codes...
 - Rewritten in C++ for v6: single data format, decoupled from lib; use internal set of interpolator routines
 - Far more maintainable; has seen speed improvements, GPU interface, extended uncertainty specs, ...
 - **Connection between global fits and GPDs? Machinery?**



13065	CT14nlo_as_0116	(barbati) (solo fit)	1	1
13066	CT14nlo_as_0117	(barbati) (solo fit)	1	1
13067	CT14nlo_as_0118	(barbati) (solo fit)	1	1
13068	CT14nlo_as_0119	(barbati) (solo fit)	1	1
13069	CT14nlo_as_0120	(barbati) (solo fit)	1	1
13070	CT14nlo_as_0121	(barbati) (solo fit)	1	1
13071	CT14nlo_as_0122	(barbati) (solo fit)	1	1
13072	CT14nlo_as_0123	(barbati) (solo fit)	1	1
13081	CT14nloIC	(barbati) (solo fit)	6	2
13090	CT14nlo_NF3	(barbati) (solo fit)	1	1
13091	CT14nlo_NF4	(barbati) (solo fit)	1	1
13092	CT14nlo_NF6	(barbati) (solo fit)	1	1
13100	CT14nlo	(barbati) (solo fit)	57	1
13158	CT14nlo_as_0111	(barbati) (solo fit)	1	2
13159	CT14nlo_as_0112	(barbati) (solo fit)	1	2
13160	CT14nlo_as_0113	(barbati) (solo fit)	1	1
13161	CT14nlo_as_0114	(barbati) (solo fit)	1	1
13162	CT14nlo_as_0115	(barbati) (solo fit)	1	1
13163	CT14nlo_as_0116	(barbati) (solo fit)	1	1
13164	CT14nlo_as_0117	(barbati) (solo fit)	1	1
13165	CT14nlo_as_0118	(barbati) (solo fit)	1	1
13166	CT14nlo_as_0119	(barbati) (solo fit)	1	1
13167	CT14nlo_as_0120	(barbati) (solo fit)	1	1

...

Event formats: from LHE to HepMC

❖ Generators

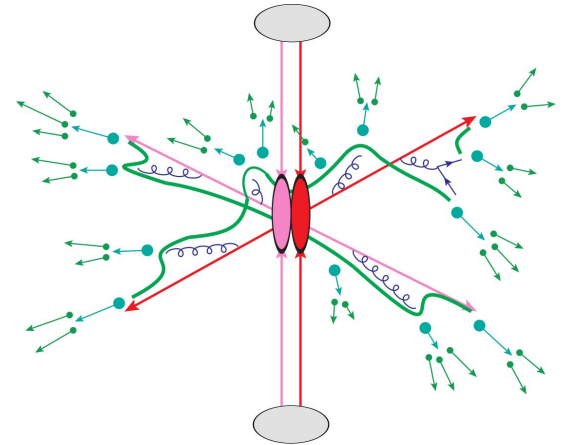
- MC codes split between ME and shower+hadronisation generators, e.g. MG5_aMC, PoWHEG, WHiZard (+ loops) vs Pythia, Herwig and Sherpa (which actually does both)
- **Loosen technical coupling via intermediate file formats:** LHE (and successor) at ME level, HepMC at hadron level.

❖ LHE / HDF5LHE

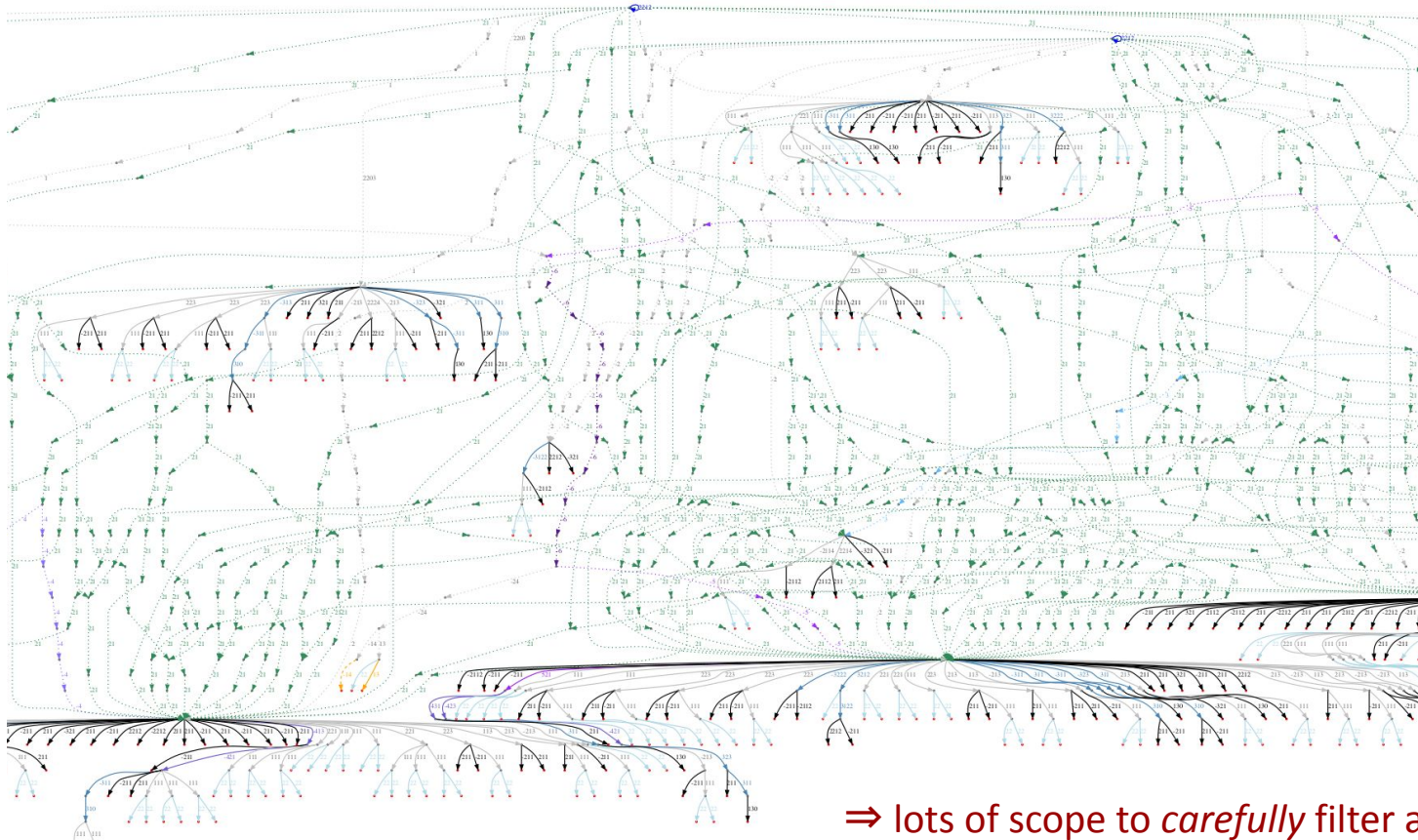
- Minimal format written for mapping LHA common blocks.
- Designed for small, usually partonic, process exchange.
- Quasi-XML → parallel-I/O HDF5 for HPC work.

❖ HepMC

- As much an [in-memory interface & data structure](#) as a file format: designed for *general* representation of events
- **NB: “truth” is not “true”!** Primarily a record of the algs



HepMC example ($\frac{1}{3}$ of a Py8 $pp \rightarrow t\bar{t}b\bar{b}$ event)



⇒ lots of scope to *carefully* filter and reduce

HepMC history and content

❖ History

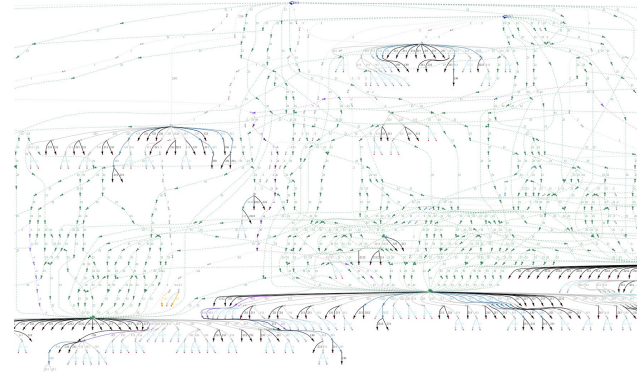
- [General graph model, of GenParticles and GenVertices](#)
- v1 in 2001: raw pointers, lack of standardisation
- v2 in 2008: resolve expt forks and CLHEP dep, adds units
- v3 in ~2019: smart ptrs, resolve constness inconsistencies, clarify ownerships, no barcodes, attributes, make memory contiguous → **ROOT-friendly**

❖ Standardisation

- v2 introduced **standardisation of particle status codes** (and later vertex statuses, but these are far less used)
- More recent [MCnet consensus on standard multiweights](#)

❖ Gaps

- **Lack of event-file manipulation tools: shouldn't have to code!**
- **Spins and polarizations still second-class. Need push to pass full e.g. spinor info, avoid guesswork**



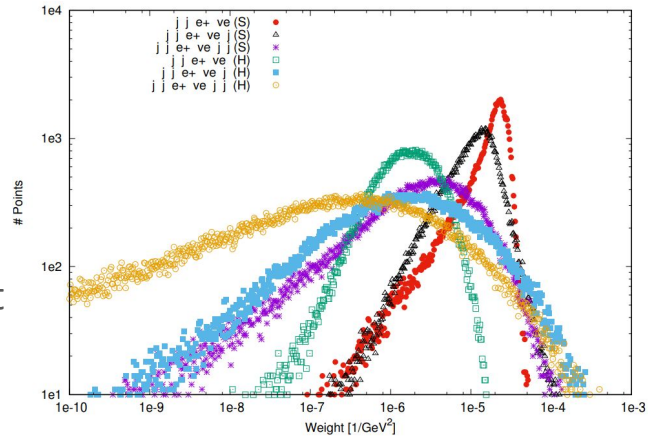
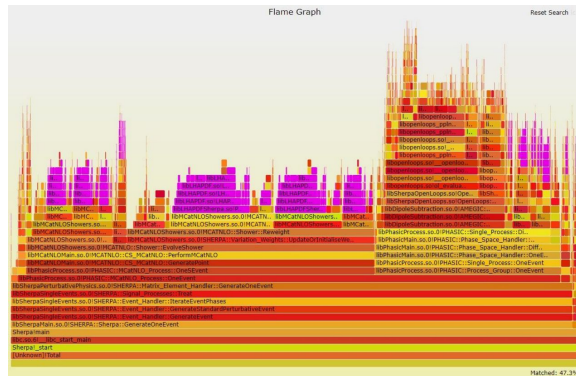
Computational performance

Bulk of LHC evgen CPU in multileg NLO V +jets, tt

- Don't forget associated packages, e.g. [LHAPDF](#), [HepMC](#)
- **40% of Sherpa CPU was in PDFs!** Joint projects to address:
 - reduce PDF calls ✓ (pilot runs, restructure code) ⇒
 - speed PDF computation ✓ (cache, vector, GPU, ipol)
 - optimise transcendental functions, specialise, ... ~

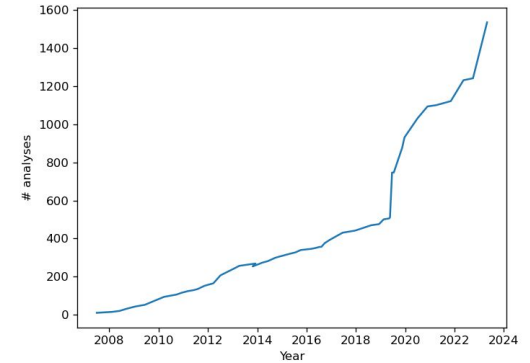
New architectures (?)

- **GPUs, other accelerators, vectorisation:** nascent efforts, not mainstream, **focus on ME since shower etc. trivially parallelisable**
- **Focus on new gen strategies using HPC**
 - **Why are multiple experiments duplicating CPU on generating the same expensive algebraic ME results?**
 - **Logistics of GPU on Grid look bad:** use systems to best match requirements ⇒ improved event formats and manipulation tools for HPC → HTC interchange



MC analysis and validation: Rivet

- ❖ **Rivet is the “LHC standard” MC analysis toolkit**
 - More broadly aiming to preserve the logic of data analyses and encourage expt-pheno collaboration
 - Availability of runnable analysis code is a game-changer!
 - Central to a *community* of analysis reinterpretation tools
- ❖ **Code-wise, a C++ core and Python tools. Some good ideas:**
 - Fiducial / **generator-independence** emphasis
 - Integration with **HepData (v4 release and new stats backend)**
 - **1500-2000 analyses**
- ❖ **Used in both MC and measurement-analysis communities: always room to grow more. Strong *network effects***
 - Requires active support in experiments, conveners who “get it”. **Lightweight reinterpretation** is accessible: powerful complement to “full” analysis preservation. Also needs stats/data: HepData, HS3, ML models, ... **plan in advance!**



Rivet analysis coverage (no searches, no heavy ion)

Rivet analyses exist for 848241 papers + 26%, 153 priority analyses required.
Total number of Inspire papers scanned = 7200, at 2025-07-02
Breakdown by identified experiment (in development):

Key	ALICE	ATLAS	CMS	LHCb	Forward	HERA	$e^+e^- (\ge 12 \text{ GeV})$	$e^+e^- (< 12 \text{ GeV})$
Rivet wanted (total):	72	111	126	183	43	461	703	647
Rivet REALLY wanted:	17	42	61	9	0	13	1	3
Rivet provided:	1436 ± 16%	136246 ± 85%	77203 ± 38%	13190 ± 7%	8151 ± 16%	8470 ± 2%	168031 ± 19%	344911 ± 38%

Show panel: [Show panel](#) [Reset panel](#)

ALICE	ATLAS	CMS	LHCb	Forward	HERA	$e^+e^- (\ge 12 \text{ GeV})$	$e^+e^- (< 12 \text{ GeV})$	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: 1502524 arXiv ID: 2006.15076 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 10
Inspire ID: 1501434 arXiv ID: 2006.09274 Report ID: CERN-EP-2020-083
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: 1796429 arXiv ID: 2004.02869 Report ID: CERN-EP-2020-025
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: 1790296 arXiv ID: 2004.03540 Report ID: CERN-EP-2020-030
Links: Inspire DOIjournal CDS arXiv HepData ATLAS_2020_11790296
- ATLAS: Measurements of the production cross-section for a Z boson in association with 3-jets in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS
Inspire ID: 1785444 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 $\sqrt{s} = 13 \text{ TeV}$

From HZTool to Rivet

- ❖ The idea of preserving experimental analyses for MC validation was born out of HZTOOL
 - HERA (H1 and ZEUS) DIS and photoproduction
 - Probing low- x , semi-perturbative physics: DIS with $Q^2 \sim 4 \text{ GeV}^2$; jet $p_T \sim 5 \text{ GeV}$; diffraction
 - Many “state of the art” models only in MCs
 - Much confusion about comparing like-with-like between generators, experiments, and analyses
 - HZTool (Fortran) for cross-experiment comparisons of similar measurements modulo cut differences

❖ Direct line to Rivet, 10 years later: “HZ mark two”

- UK e-science funding; adopted by EU MCnet network



Aim: Study of future physics potentials at HERA in collider and fixed target modes, including high luminosity, polarized beams and nuclei.

[Proceedings of the Workshop](#)

[Old home page](#) and [workshop meetings](#)



Working Groups:

- Structure Functions
- Electroweak Physics
- Beyond the Standard Model
- Heavy Quark Production and Decay
- Jets and High p_T Phenomena
- Diffractive Hard Scattering
- Polarized Protons and Electrons
- Light and Heavy Nuclei in HERA
- HERA Upgrades and Impacts on Experiments



Organizing Committee:
Gunnar Ingelman, Uppsala/DESY (Chairman)
Albert De Roeck, DESY
Robert Klanner, DESY



Secretary:
Ms. H. Haertel
DESY-FH1K
Notkestrasse 85
D-22603 Hamburg
Phone: +49-40-8998-3105
Fax: +49-40-8998-3093

Email: heraus96@mail.desy.de

Advisory Committee:
W.Buchmüller, J.Feltesse, A.Levy,
H.Schröder, J.van den Brand, A.Wagner

If you are using mosaic, click [here](#).

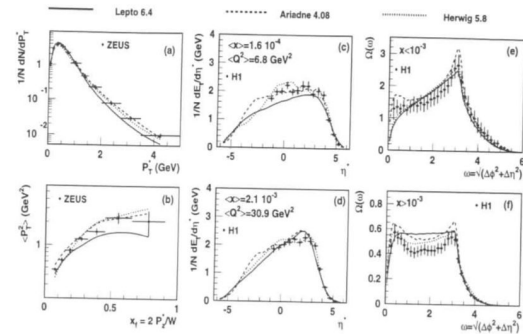


Figure 1: The transverse momenta dN/dp_T (a) and the ‘seagull’ plot $\langle P_T^2 \rangle \times x_F$ (b) of single particles in the positive hemisphere of the hadronic center of mass. The transverse energy flow $dE_T/d\eta$ in a low (c) and high (d) x and Q^2 bin. The transverse energy-energy correlations for $x > 10^{-3}$ (e) and $x < 10^{-3}$ (f).

Lessons learned

❖ A simple/obvious idea, with surprising impact:

- **Reproducing a key plot (or not) is *powerful***

⇒ *understand physics, communicate issues, improve MCs*

- **A common language for phenomenology and experiment**

❖ But...

- “Obvious” to use partons, bosons, etc. direct from the event graph

- Frequently unphysical, depend on approximations. May not even exist!

- Scalability of many analyses to new MCs means avoiding gen-dependence

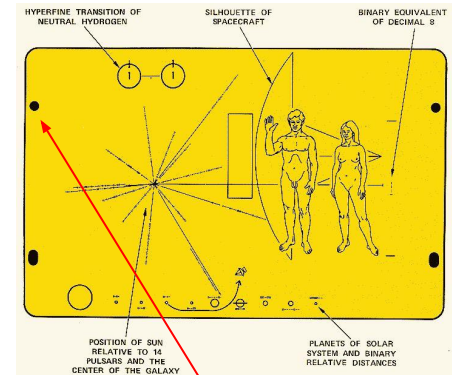
⇒ **predict “real” observables, from well-defined final states**

❖ Standardisation: boring but important

- (physical) event format conventions, statuses, PDG particle numbering, weights...

❖ Scalability

- Lots of expensive operations are repeated: sharing calculations is essential



Physically safe analysis methods



Avoiding unstandardised event-graph features was pragmatic, but led to some genuine physical insights:

- ❖ **Refining the “fiducial analysis”**, define *unfolding targets*

- ❖ **Hadronisation as a “decoherence barrier”**

use the natural dividing line between the quantum-interfering hard process & semi-classical decays: ~ no tempting partons!

- ❖ **Bringing truth tagging closer to reco**

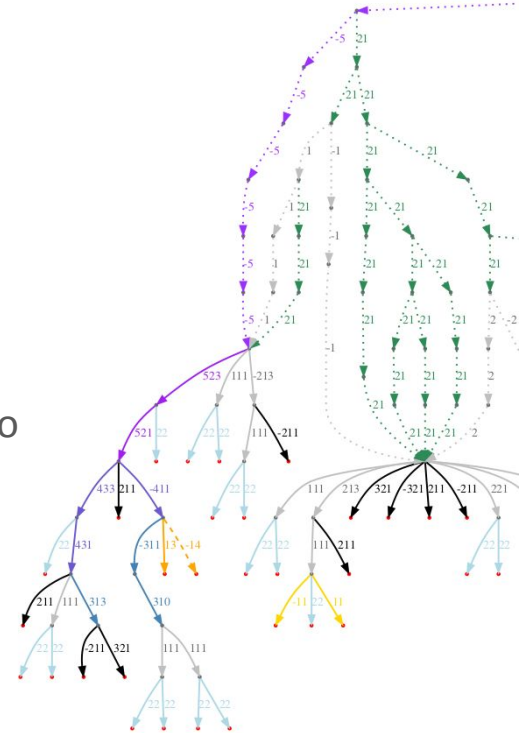
first releases used *b*-ancestry of jet constituents to set HF labels: too inclusive! \Rightarrow associate the hard-fragmenting, weakly-decaying *B*

- ❖ **Promptness/directness tests**

don't identify a particle “from the hard process”; do it backward. Label as *indirect* via recursive checks for hadron parentage

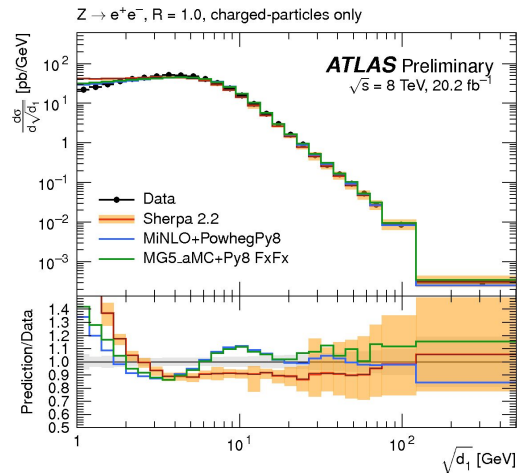
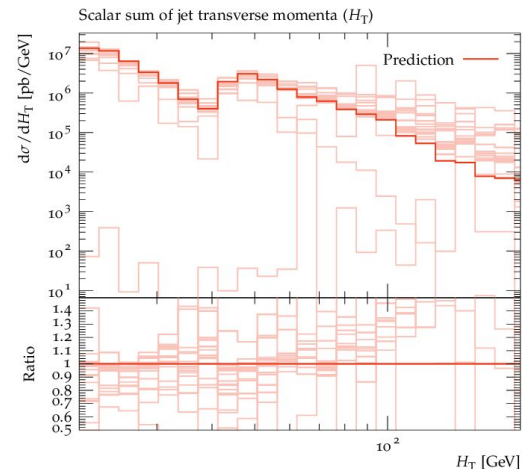
- ❖ **Dressed leptons**

we now primarily *dress* truth leptons with their photon halo



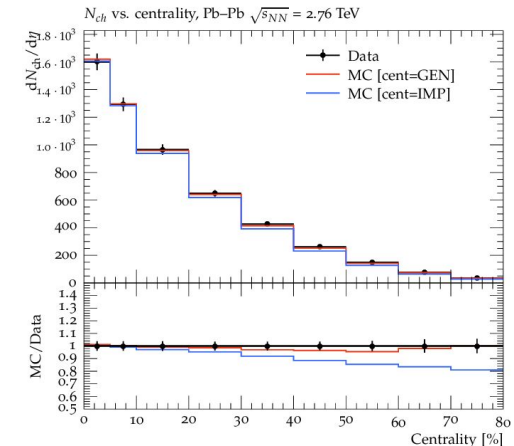
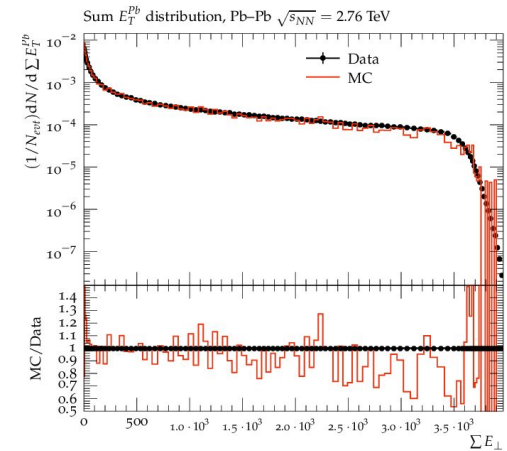
Multiweights and re-entry

- ❖ MC weight vectors allow expression of increasingly complex theory uncertainties. But a burden for analysis chains: have to propagate and correctly combine $O(200)$ weight streams!
- ❖ Rivet 3: complex automatic handling of weights
~invisible to users: data objects *look* like histograms etc. but are secretly multiplexed
- ❖ Can now re-call finalisation to combine runs:
RAW histogram stage preserves pre-finalize objects
⇒ “re-entrant” perfect rivet-merge-ing
Key for e.g. pA/pp or W/Z ratios, + BSM recasting
- ❖ Data types are important: glimpses of a fully coherent separation of semantics from presentation



Heavy-ion physics preservation

- ❖ “Adding heavy-ion support” sounds trivial!
- ❖ Actually nuanced \Rightarrow lots of structural impacts
 - HI observables often require centrality-fraction calibration curves: we need a 2-pass run.
 - Flow observables, event/event correlations... all centrality-binned!
 - Swappable definitions: few HI generators are general-purpose enough to do “everything”
- ❖ All supported “out of the box” since v3
 - Paper: <https://arxiv.org/abs/2001.10737>
 - Core development tool for Pythia/Angantyr: authors and ALICE (etc.) collaborators providing analyses
- ❖ Should be well-placed for EIC HI requirements



HI community engagement

- ❖ Great “spontaneous” engagement both within BNL HI. Several productive workshops

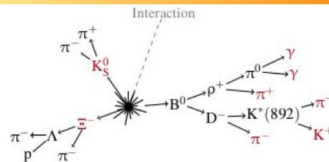
Summary

1. Data getting into HEPData

2. Build your own undergraduate army



3. Primary particle definition



4. Validation Procedure

5. HEPMC output may have some issues

<https://indico.bnl.gov/event/10966/>

- ❖ HepData, Rivet
- ❖ Better ex/ph communication
- ❖ Faster model/data comparisons
- ❖ Addressing issues with formats and incomplete models
- ❖ Undergrad army!

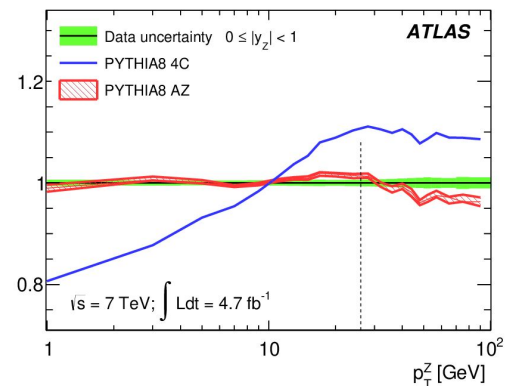
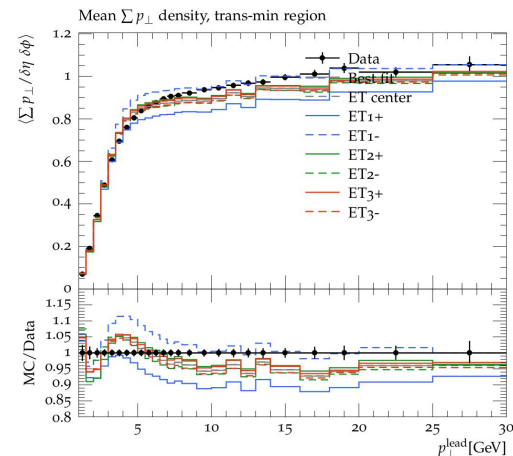
Christine Nattrass, University of Tennessee, Knoxville, HF-QGP Rivet 2021

Work reactivated recently, very relevant overlap with EIC community

Applications: from tuning to...

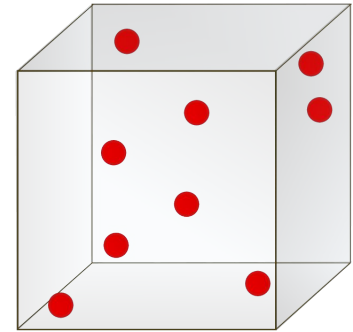
First “killer app”: huge pre-LHC soft-QCD uncertainties:

- ❖ Tuning required Rivet analyses from expt
- ❖ Feed in to underlying event, pile-up, etc. modelling
 - Better tunes \Rightarrow better analysis, better results
 - **Impact:** LEP and Tevatron analyses published for ~ 10 years suddenly got *used!* And cited...
 - \Rightarrow ATLAS tunes, CMS tunes, **eigentunes...**
 - \Rightarrow Rapid responses to preliminary data
 - **Model development:** matching & merging, addition of MPI energy-evolution & colour-reconnection to Herwig, ...
- ❖ More recently, use of Rivet’s analysis collection for *BSM, Higgs EFT, PDF fits (see Anjelo’s talk after), ...*
 - **Good analysis preservation enables new ideas**



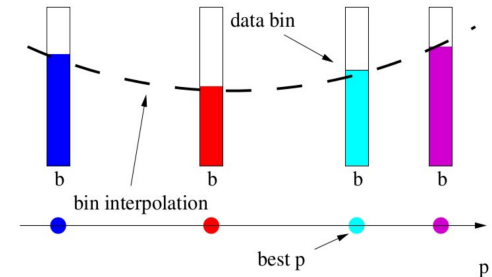
Speaking of tuning... Professor/Apprentice

- ❖ MC models contain $O(10-30)$ influential free parameters, typically in non-perturbative parts
 - Fragmentation kinematics, cluster and string splittings, flavour and meson/baryon mixtures, MPI, shower scales
 - No *a priori* predictions \Rightarrow “tune” to match data
 - As much data as possible — within model capabilities



- ❖ [Professor](#) (2009) and [Apprentice](#) (2021) semi-automate the fit, with a surrogate-model approach

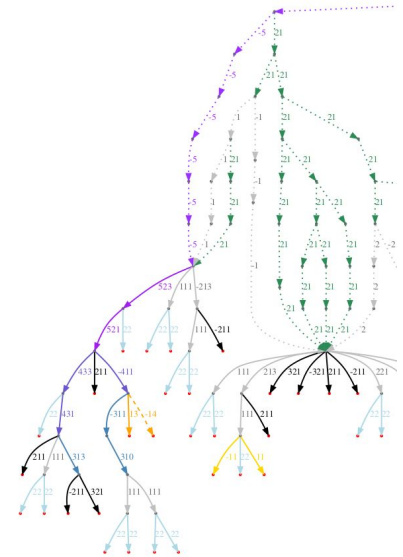
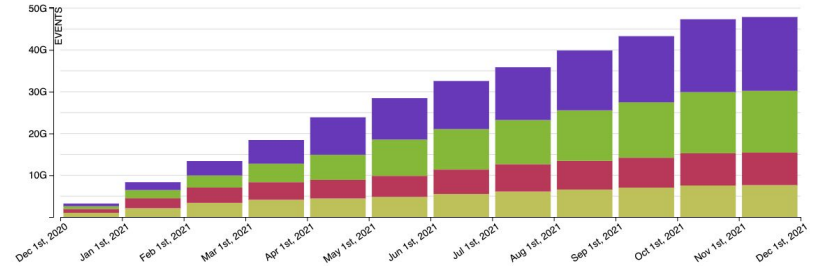
- pre-sample from the MC parameter (sub-)space
- parallel-execute a comprehensive set of MC runs with (Rivet) analysis \rightarrow histogram predictions vs data
- parametrise the response of *each bin* to the parameters
- \Rightarrow usual iterative fit to reference data



- ❖ Supported & developed... but **currently** at “best effort” level

LHC experience of MC production logistics

- ❖ Both ATLAS and CMS batch MC event generation in *campaigns* of $O(50 \text{ Gevt})$
 - not every year: often reuse+extend evgen
- ❖ Running generators at scale introduces exciting new failure modes!
 - rare numerical issues
 - **configuration mistakes very costly**
 - **requires serious software and configuration management**
- ❖ MC generation has particularly strong *upward* CPU scaling
 - analysts have accustomed to MC being a good proxy for data.
 - **Always demand for next order in precision**
 - ⇒ typically costs an *order of magnitude* more CPU!
 - sometimes the right answer is “no”... match precision to requirements! [While being impressed it's *possible*]



Code management and integration

❖ Code and build management

- LHC experiment software mostly built on LCG software bundles
- Generators included via GenSer project: integrate release tarballs, perform basic tests
- Used directly by ATLAS and LHCb, tarballs by CMS, not by ALICE
- Experiment frameworks need “glue” packages to pick up compiled generators. Library version compatibility, e.g. FastJet & HepMC, not always coherent...
- Not 100% clear the intermediate has been worth it: generators are out-of-the-box packaged following far more usable software methods than in the Fortran era

❖ Experiment-framework interfacing

- Dedicated packages for each generator — configurations *still* not all programmatically friendly (Powheg, MG5). Do as a little physics as possible in interfaces!
- Documentation: prefer repo READMEs to wikis!!
- Post-proc event-graph testing and fixing: flag common problems like unknown PIDs, broken graphs, unexpected displaced vertices, E & p imbalances...

Sample-configuration management

❖ Python job options (JOs) in both ATLAS and CMS

- $O(10,000-1,000,000)$ sample configurations: jet slices, heavy-flavour filtering, BSM param-scans, other enhancement-biasing...
- **New ones will be made by non-experts, via copy & paste** \Rightarrow validation essential!
- Vigilance needed to identify common elements and manage common JO snippets. Chain snippets for e.g. standard params, tunes, modes, ...



❖ Managing sample requests & production status

- “Keep it simple. I'm sure a GitLab issue would work just fine”. GitLab is great...
- CMS more sophisticated than ATLAS: dedicated Web apps, e.g. GrASP, vs Twiki+GSheets+JIRA+...
- Not “interesting”, not physics, but needs effort and dev time

❖ Distribution

- JO updates far too frequent to include in sw releases. Sync via CVMFS or tarballs. Need *versioning*: configs need to be exactly repeatable, for sample extensions

The screenshot shows the GrASP web interface. At the top, it says 'GrASP' and 'Logged in as: Justinas Rumsevicius'. Below this, there are two main sections: 'Existing Samples' and 'Future Campaign Planning'. Each section contains a table with columns for Campaign Name and various physics group abbreviations (B2G, BPH, BTM, EGM, EXO, HCA, HGC, HIG, HIN, JME, LUM, MUO, PPS, SMP, SUS, TAU, TOP, TRK, TSG). The 'Existing Samples' table lists three campaigns from Summer 2016 to 2018. The 'Future Campaign Planning' table lists three campaigns for Summer 190L16. Below these tables, there is a 'User Tagged Samples' section with a 'Tag' input field, a 'Wmass_Samples' button, and an 'Add new tag' button.

Physics content & communication

❖ Experiment generator experts rely on MC authors

- Get direct bug/task reporting with authors
- Need regular interaction, understanding of different motivations and incentives (theory vs exp worlds)
- Need ability to supply standalone configs to authors: they can't do anything with JOs ⇒ design interfaces with ability to dump standalone steering files
- don't try to be too clever!

❖ Standardising

- Work with authors to standardise: formats, systematics weight structure, PDG codes
- evolving and enforcing standards makes everything better/clearer in the long run

Contents

1	Introduction	2
2	Software and Computing Challenges	5
3	Programme of Work	11
3.1	Physics Generators	11
3.2	Detector Simulation	15
3.3	Software Trigger and Event Reconstruction	23
3.4	Data Analysis and Interpretation	27
3.5	Machine Learning	31
3.6	Data Organisation, Management and Access	36
3.7	Facilities and Distributed Computing	41
3.8	Data-Flow Processing Framework	44
3.9	Conditions Data	47
3.10	Visualisation	50
3.11	Software Development, Deployment, Validation and Verification	53
3.12	Data and Software Preservation	57
3.13	Security	60
4	Training and Careers	65
4.1	Training Challenges	65
4.2	Possible Directions for Training	66
4.3	Career Support and Recognition	68
5	Conclusions	68

Summary / thoughts

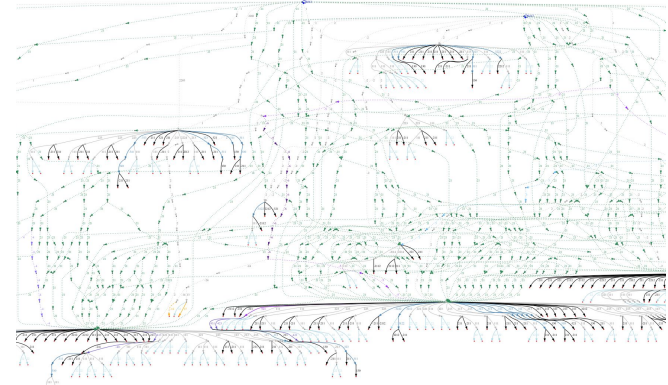
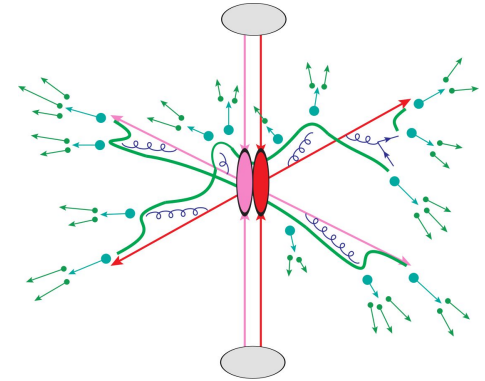
- ❖ **We have a well-developed toolchain of supporting codes around the MC generators themselves.**
 - From MCnet historically, schemes like UK SWIFT-HEP helping: more please! Note [MC tools workshop here at the end of June](#)
 - Most tools like LHE, HepMC, LHAPDF, Rivet, Prof... not funded!
- ❖ **As well as production tools, analysis preservation matters**
 - Completes the virtuous cycle from theory to exp and back
 - **Rivet+HepData** are ready for EIC: key is incorporation of preservation in experiment procedures and incentives
 - Reusability enhances analysis/publication impact, and provides fruitful exp/th collaboration, esp for junior scientists
- ❖ **Lots of LHC-era experience in production scaling and issues**
 - Mistakes are costly! Let's avoid having to re-learn the hard way.
LHC-EIC communication routes? Tool re-use? Investment in process



Backup slides

MC generation

- ❖ **MC generation is where theory meets experiment**
 - The fundamental pp , pA , AA collision, *sans* detector
- ❖ **Components of an “exclusive” event-generator chain:**
 - QFT **matrix element** sampling at fixed-order in QCD
 - *Dressed* with approximate collinear splitting functions, iterated in factorised Markov-chain “**parton showers**”
 - FS parton evolution terminated at $Q \sim 1$ GeV: phenomenological **hadronisation** modelling
 - Mixed with **multiple partonic interaction** modelling
 - Finally particle **decays**, and other niceties
- ❖ **Modern HEP is hostage to shower MCs!**
 - The main mechanism for translating theory to experimental signatures, from QCD to BSM
 - Generally very complex modelling and output



Generator balance

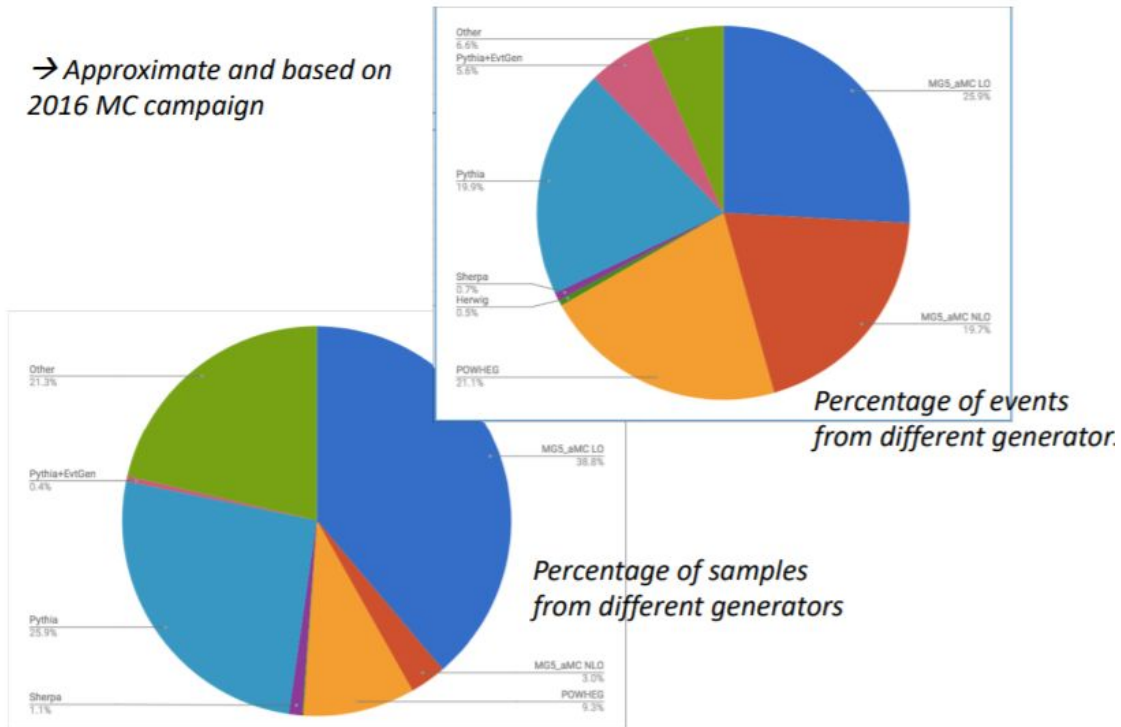
❖ CMS has a particularly strong reliance on MG5_aMC and Pythia

- 97% Pythia showering!
- few percent Herwig and Sherpa variations

❖ ATLAS a much broader set of generators:

- Sherpa NLO for V+jets, and VV processes
- Powheg+Pythia for $t\bar{t}$
- MG5 LO for most BSM

→ Approximate and based on 2016 MC campaign



from Efe Yazgan, CMS

Generation practicalities

❖ Logistics of bulk NLO production

- With “generator generators”, job splitting to $O(1\text{kevt})$ → don’t waste time rebuilding the model and remapping the phase-space (integration) in each job
- Prebuild “gridpacks” storing results of integration: produced privately on HPC.
Gridpack distribution with JOs / via CVMFS
- LHE generation and tracking correspondence to showered HepMC event (esp. with post-shower event-filtering). Embed in HepMC3?

❖ Weighting, filtering and enhancement

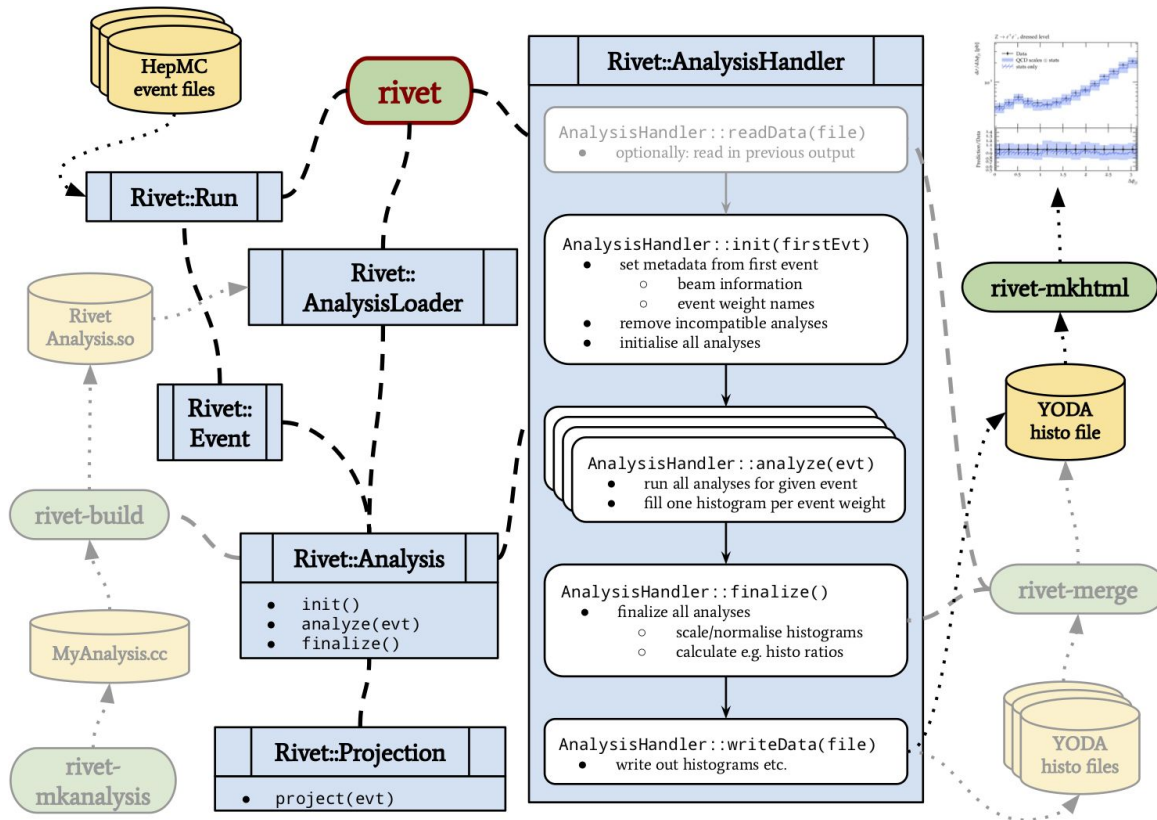
- Systematics weights: currently $O(100)$ for ME scales, PDFs, sometimes shower vars
- Post-hoc filtering: focus samples on flavour combinations & phase-space of interest
- Increasingly “enhancing” phase-space coverage with biased sampling and counter-weights: adaptive samplers in multileg codes “learn” biases, so efficient

❖ Data formats

- Automatic persistency is overrated — if any risk of change, use dedicated TP converters to persist key event info and handle compatibility. Annoying, but...
- Downstream analysis formats: reduce event graph to collections of standard e.g. (many different) truth jets with truth flavour-tags dressed leptons, truth MET, etc.

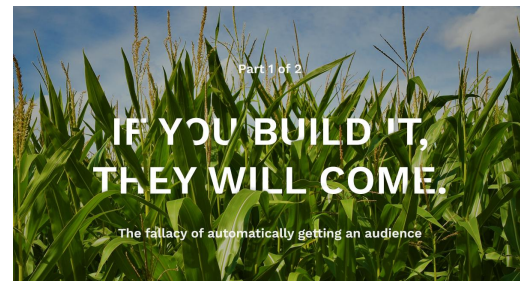
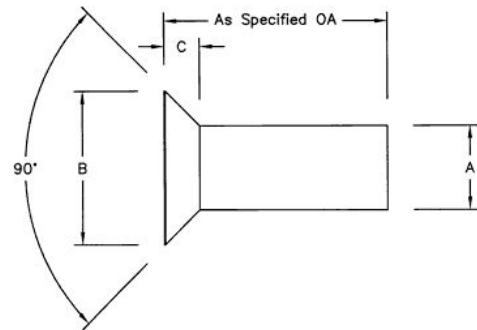
The result

- ❖ Rivet v3 structure [arXiv:1912.05451](https://arxiv.org/abs/1912.05451)
- ❖ Streamlined set of tools from analysis coding to event processing to plotting (and other applications)
- ❖ And a key gateway to connect data analysis to theory (and back again)



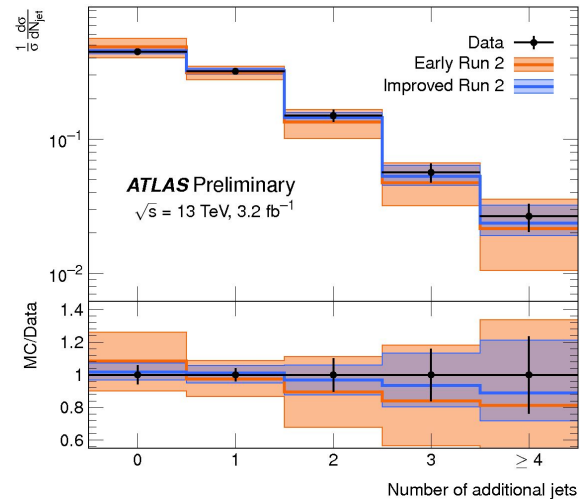
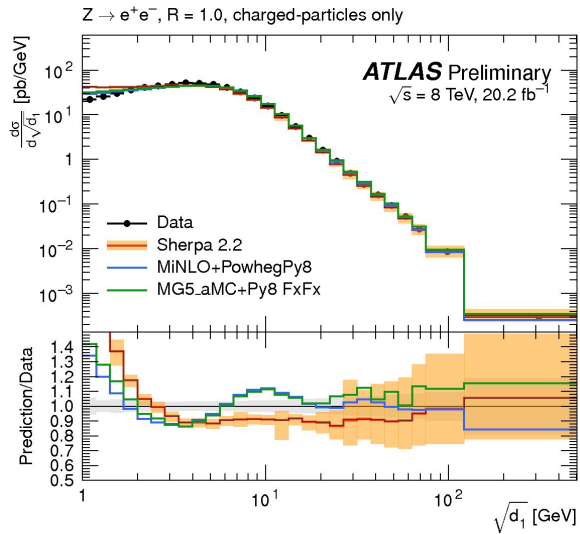
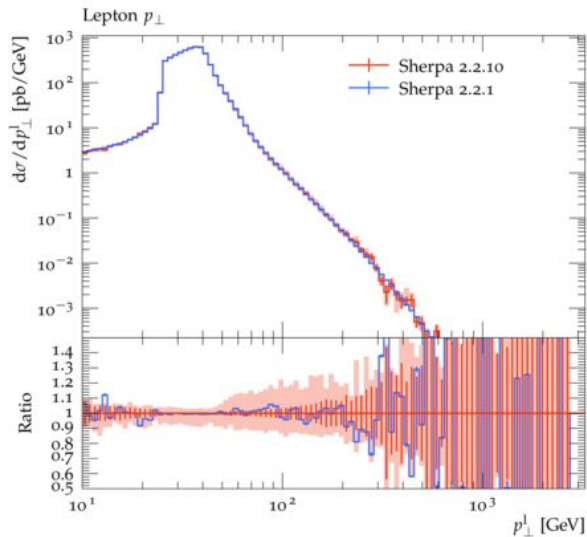
Designing Rivet

- ❖ **Ease of use**
 - **Big emphasis on “more physics, less noise”!**
 - Minimal boilerplate analysis code, HepData sync
 - Event loop and histogramming basically familiar
 - **Tools to avoid having to touch the raw event graph**
- ❖ **Embeddable**
 - OO C++ library, Python wrapper, sane user scripts
 - Generator independence: communication via HepMC
 - Note HepMC3 HI-support efforts
 - Analysis routines factorised: loaded as “plugins”
- ❖ **Efficient**
 - **Avoid recomputations via “projection” caching system**
- ❖ **Physical**
 - **Measurements primarily from final-state particles only**



MC systematics bands via multiweights

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



Weight-naming standardisation: see [arXiv:2203.08230](https://arxiv.org/abs/2203.08230)

Getting & using Rivet

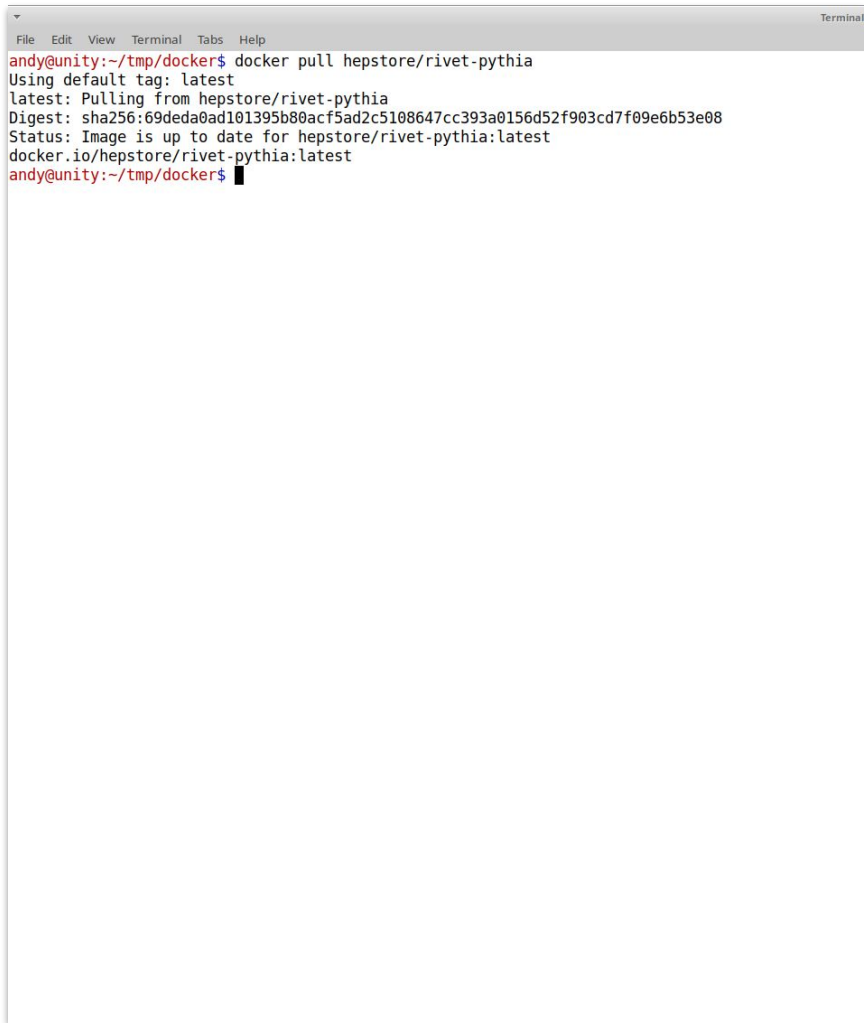
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ [copy an existing analysis!](#)

A terminal window with a menu bar (File, Edit, View, Terminal, Tabs, Help) and a title bar (Terminal). The terminal shows the following text:

```
andy@unity:~/tmp/docker$ docker pull hepstore/rivet-pythia
Using default tag: latest
latest: Pulling from hepstore/rivet-pythia
Digest: sha256:69deda0ad101395b80acf5ad2c5108647cc393a0156d52f903cd7f09e6b53e08
Status: Image is up to date for hepstore/rivet-pythia:latest
docker.io/hepstore/rivet-pythia:latest
andy@unity:~/tmp/docker$
```

Getting & using Rivet

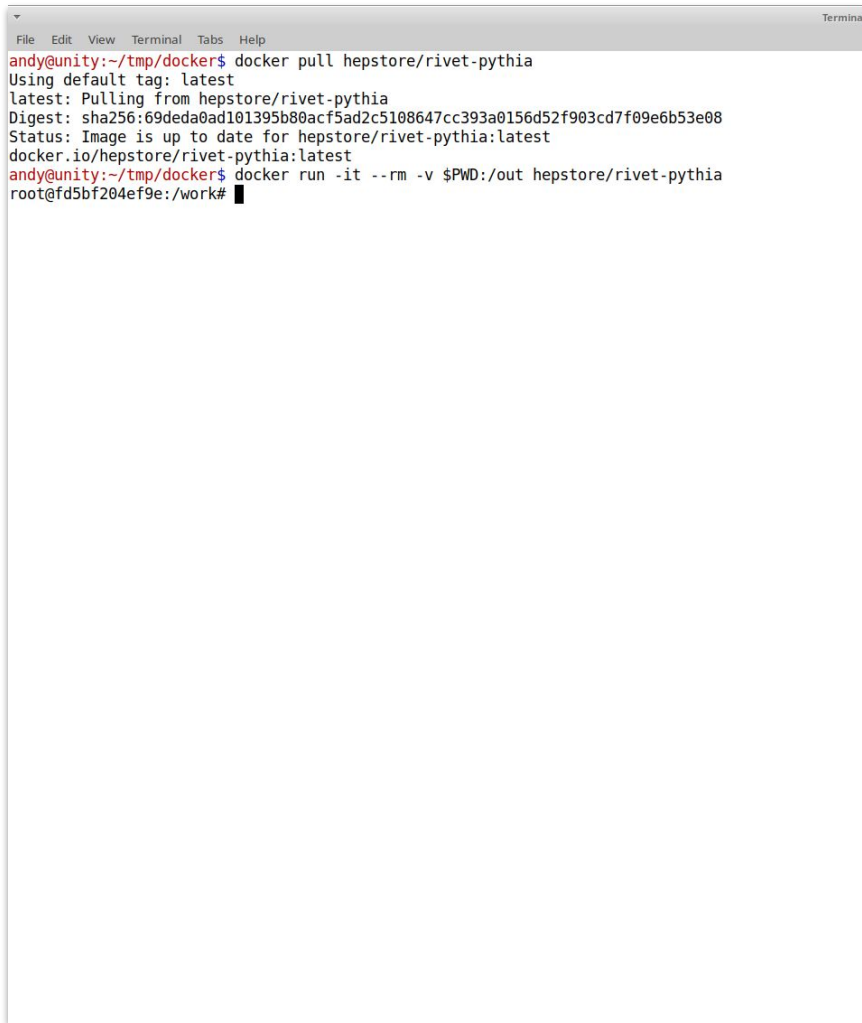
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



```
File Edit View Terminal Tabs Help
andy@unity:~/tmp/docker$ docker pull hepstore/rivet-pythia
Using default tag: latest
latest: Pulling from hepstore/rivet-pythia
Digest: sha256:69deda0ad101395b80acf5ad2c5108647cc393a0156d52f903cd7f09e6b53e08
Status: Image is up to date for hepstore/rivet-pythia:latest
docker.io/hepstore/rivet-pythia:latest
andy@unity:~/tmp/docker$ docker run -it --rm -v $PWD:/out hepstore/rivet-pythia
root@fd5bf204ef9e:/work#
```

Getting & using Rivet

**Lightweight analysis preservation
is valuable... and easy to start**

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

A terminal window with a grey title bar containing the word "Terminal". The terminal content shows two lines of shell commands and their outputs. The first line is a copy command: `root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd`. The second line is a command to open a file in nano: `root@d8c06acf8f66:/work# nano py.cmnd`. The cursor is at the end of the second line.

```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
```

Getting & using Rivet

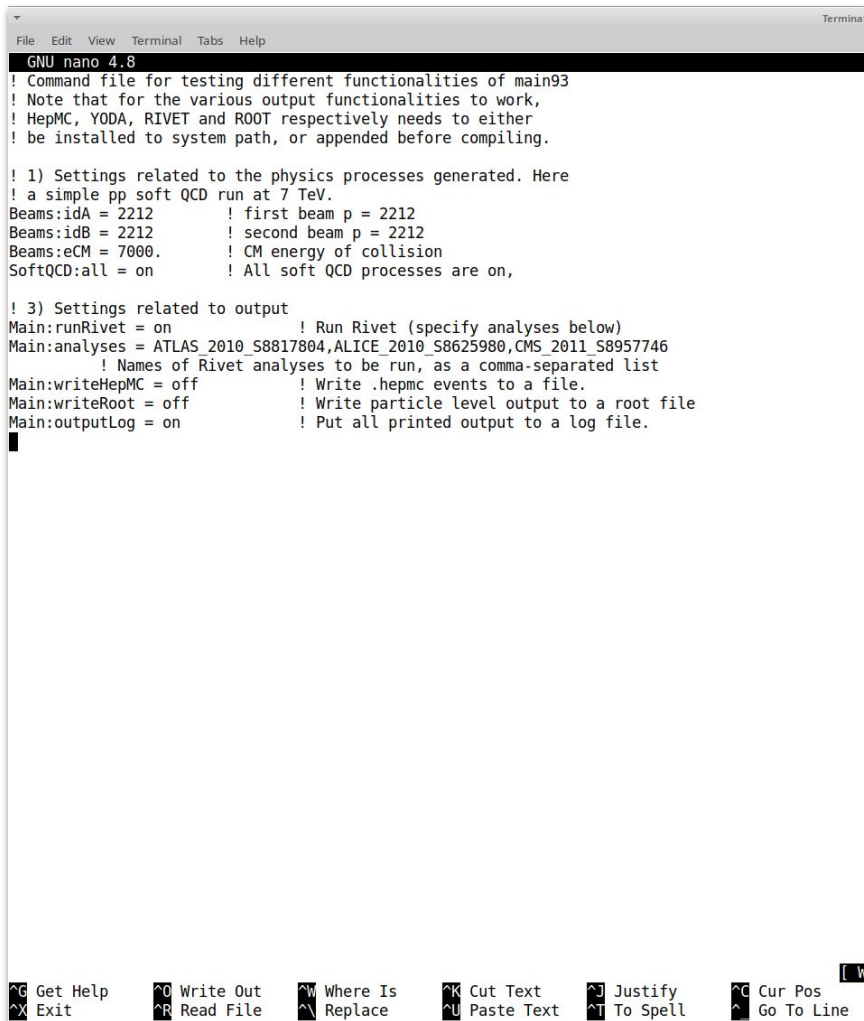
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



```
GNU nano 4.8
! Command file for testing different functionalities of main93
! Note that for the various output functionalities to work,
! HepMC, YODA, RIVET and ROOT respectively needs to either
! be installed to system path, or appended before compiling.

! 1) Settings related to the physics processes generated. Here
! a simple pp soft QCD run at 7 TeV.
Beams:idA = 2212      ! first beam p = 2212
Beams:idB = 2212      ! second beam p = 2212
Beams:eCM = 7000.     ! CM energy of collision
SoftQCD:all = on      ! All soft QCD processes are on,

! 3) Settings related to output
Main:runRivet = on      ! Run Rivet (specify analyses below)
Main:analyses = ATLAS_2010_S8817804,ALICE_2010_S8625980,CMS_2011_S8957746
! Names of Rivet analyses to be run, as a comma-separated list
Main:writeHepMC = off  ! Write .hepmc events to a file.
Main:writeRoot = off   ! Write particle level output to a root file
Main:outputLog = on    ! Put all printed output to a log file.
█

^G Get Help      ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify     ^C Cur Pos
^X Exit          ^R Read File    ^N Replace      ^U Paste Text   ^T To Spell    ^_ Go To Line
```

Getting & using Rivet

**Lightweight analysis preservation
is valuable... and easy to start**

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000
```


Getting & using Rivet

**Lightweight analysis preservation
is valuable... and easy to start**

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000
```

Getting & using Rivet

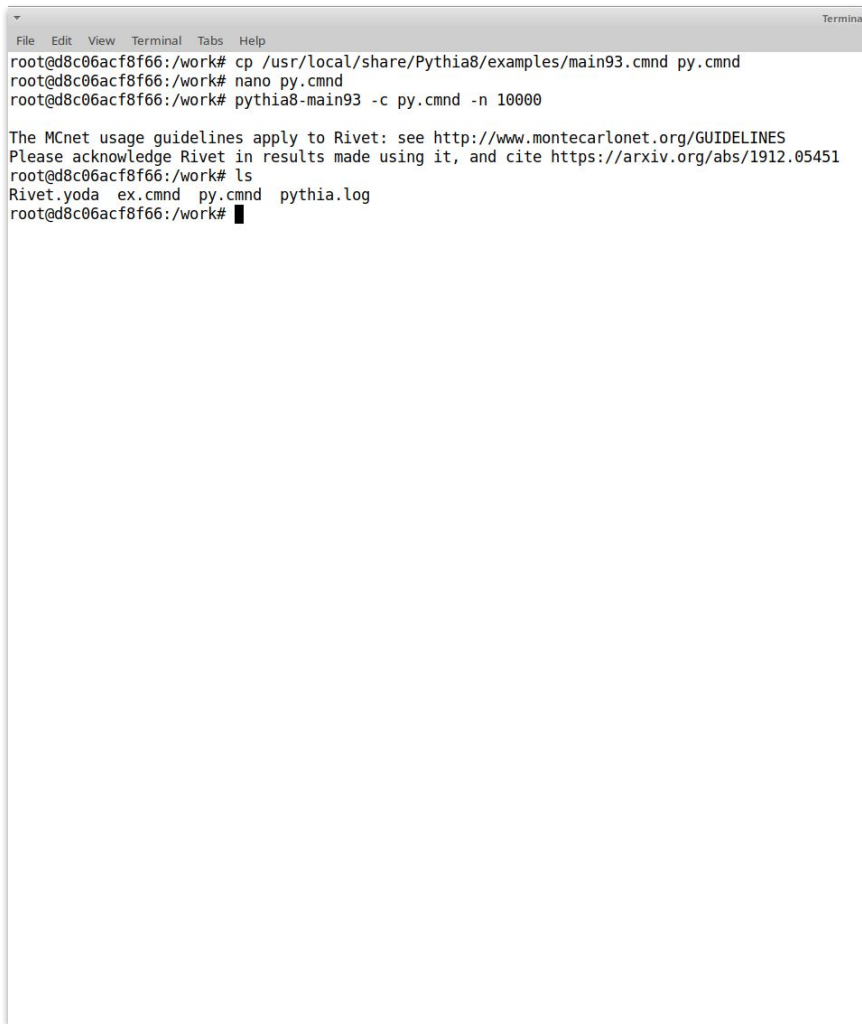
Lightweight analysis preservation is valuable... and easy to start

As either a “user” or analysis author, the barrier is lower than ever: we recommend using our **Docker images** to get started

Ideal for student projects!

Tutorials available from the [Rivet website](#), a **walkthrough** in the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

A terminal window showing a sequence of commands and their outputs. The user copies a file, creates a new file with nano, and runs a command to execute a script. The output includes a notice about MCnet usage guidelines and a list of files in the current directory.

```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000

The MCnet usage guidelines apply to Rivet: see http://www.montecarlonet.org/GUIDELINES
Please acknowledge Rivet in results made using it, and cite https://arxiv.org/abs/1912.05451
root@d8c06acf8f66:/work# ls
Rivet.yoda  ex.cmnd  py.cmnd  pythia.log
root@d8c06acf8f66:/work#
```

Getting & using Rivet

Lightweight analysis preservation is valuable... and easy to start

As either a “user” or analysis author, the barrier is lower than ever: we recommend using our **Docker images** to get started

Ideal for student projects!

Tutorials available from the [Rivet website](#), a **walkthrough** in the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000

The MCnet usage guidelines apply to Rivet: see http://www.montecarlo.net.org/GUIDELINES
Please acknowledge Rivet in results made using it, and cite https://arxiv.org/abs/1912.05451
root@d8c06acf8f66:/work# ls
Rivet.yoda  ex.cmnd  py.cmnd  pythia.log
root@d8c06acf8f66:/work# rivet-mkhtml Rivet.yoda
```

Getting & using Rivet

Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

```
File Edit View Terminal Tabs Help
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000

The MCnet usage guidelines apply to Rivet: see http://www.montecarlo.net.org/GUIDELINES
Please acknowledge Rivet in results made using it, and cite https://arxiv.org/abs/1912.05451
root@d8c06acf8f66:/work# ls
Rivet.yoda ex.cmnd py.cmnd pythia.log
root@d8c06acf8f66:/work# rivet-mkhtml Rivet.yoda
Making 35 plots
Plotting ./rivet-plots/ALICE_2010_S8625980/Nevt_after_cuts.dat (35/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d03-x01-y01.dat (34/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d06-x01-y01.dat (33/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d01-x01-y01.dat (32/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d02-x01-y01.dat (31/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d03-x01-y01.dat (30/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d04-x01-y01.dat (29/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d05-x01-y01.dat (28/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d06-x01-y01.dat (27/35 remaining)
█
```

Getting & using Rivet

Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our

Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

```
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000

The MCnet usage guidelines apply to Rivet: see http://www.montecarlo.net.org/GUIDELINES
Please acknowledge Rivet in results made using it, and cite https://arxiv.org/abs/1912.05451
root@d8c06acf8f66:/work# ls
Rivet.yoda ex.cmnd py.cmnd pythia.log
root@d8c06acf8f66:/work# rivet-mkhtml Rivet.yoda
Making 35 plots
Plotting ./rivet-plots/ALICE_2010_S8625980/Nevt_after_cuts.dat (35/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d03-x01-y01.dat (34/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d06-x01-y01.dat (33/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d01-x01-y01.dat (32/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d02-x01-y01.dat (31/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d03-x01-y01.dat (30/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d04-x01-y01.dat (29/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d05-x01-y01.dat (28/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d06-x01-y01.dat (27/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d07-x01-y01.dat (26/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d08-x01-y01.dat (25/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d09-x01-y01.dat (24/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d10-x01-y01.dat (23/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d11-x01-y01.dat (22/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d12-x01-y01.dat (21/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d13-x01-y01.dat (20/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d14-x01-y01.dat (19/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d15-x01-y01.dat (18/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d16-x01-y01.dat (17/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d17-x01-y01.dat (16/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d18-x01-y01.dat (15/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d19-x01-y01.dat (14/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d20-x01-y01.dat (13/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d21-x01-y01.dat (12/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d22-x01-y01.dat (11/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d23-x01-y01.dat (10/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d24-x01-y01.dat (9/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d25-x01-y01.dat (8/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d26-x01-y01.dat (7/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d01-x01-y01.dat (6/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d02-x01-y01.dat (5/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d03-x01-y01.dat (4/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d04-x01-y01.dat (3/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d05-x01-y01.dat (2/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d06-x01-y01.dat (1/35 remaining)
root@d8c06acf8f66:/work#
```


Getting & using Rivet

Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our

Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

```
root@d8c06acf8f66:/work# cp /usr/local/share/Pythia8/examples/main93.cmnd py.cmnd
root@d8c06acf8f66:/work# nano py.cmnd
root@d8c06acf8f66:/work# pythia8-main93 -c py.cmnd -n 10000

The MCnet usage guidelines apply to Rivet: see http://www.montecarlo.net.org/GUIDELINES
Please acknowledge Rivet in results made using it, and cite https://arxiv.org/abs/1912.05451
root@d8c06acf8f66:/work# ls
Rivet.yoda ex.cmnd py.cmnd pythia.log
root@d8c06acf8f66:/work# rivet-mkhtml Rivet.yoda
Making 35 plots
Plotting ./rivet-plots/ALICE_2010_S8625980/Nevt_after_cuts.dat (35/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d03-x01-y01.dat (34/35 remaining)
Plotting ./rivet-plots/ALICE_2010_S8625980/d06-x01-y01.dat (33/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d01-x01-y01.dat (32/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d02-x01-y01.dat (31/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d03-x01-y01.dat (30/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d04-x01-y01.dat (29/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d05-x01-y01.dat (28/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d06-x01-y01.dat (27/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d07-x01-y01.dat (26/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d08-x01-y01.dat (25/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d09-x01-y01.dat (24/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d10-x01-y01.dat (23/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d11-x01-y01.dat (22/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d12-x01-y01.dat (21/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d13-x01-y01.dat (20/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d14-x01-y01.dat (19/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d15-x01-y01.dat (18/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d16-x01-y01.dat (17/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d17-x01-y01.dat (16/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d18-x01-y01.dat (15/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d19-x01-y01.dat (14/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d20-x01-y01.dat (13/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d21-x01-y01.dat (12/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d22-x01-y01.dat (11/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d23-x01-y01.dat (10/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d24-x01-y01.dat (9/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d25-x01-y01.dat (8/35 remaining)
Plotting ./rivet-plots/ATLAS_2010_S8817804/d26-x01-y01.dat (7/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d01-x01-y01.dat (6/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d02-x01-y01.dat (5/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d03-x01-y01.dat (4/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d04-x01-y01.dat (3/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d05-x01-y01.dat (2/35 remaining)
Plotting ./rivet-plots/CMS_2011_S8957746/d06-x01-y01.dat (1/35 remaining)
root@d8c06acf8f66:/work# cp -r rivet-plots/ /out/
root@d8c06acf8f66:/work# █
```

Getting & using Rivet

Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

A terminal window titled "Terminal" with a menu bar (File, Edit, View, Terminal, Tabs, Help) and a title bar (root@d8c06acf8f66: /work). The terminal shows a user named "andy" at a shell prompt in a Docker container. The user enters "ls" and the output is "rivet-plots". The prompt then changes to "andy@unity:~/tmp/docker\$".

```
root@d8c06acf8f66: /work
andy@unity:~/tmp/docker$ ls
rivet-plots
andy@unity:~/tmp/docker$
```

Getting & using Rivet

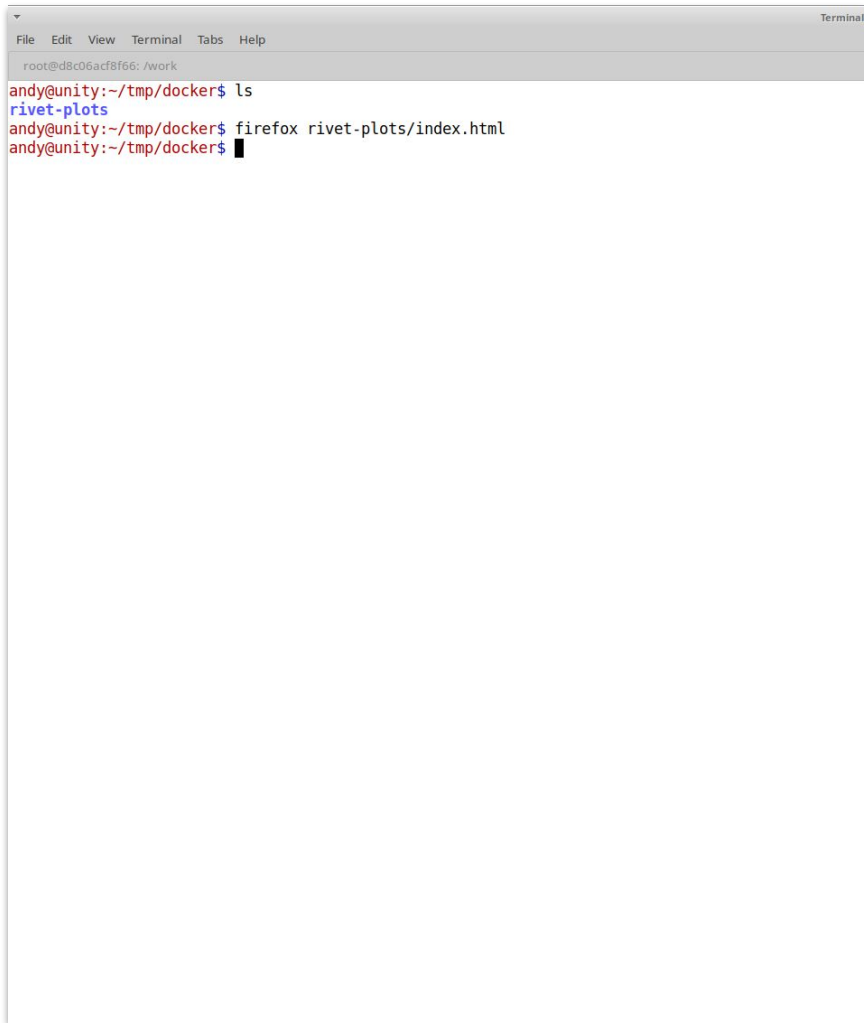
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!

A terminal window titled "Terminal" with a menu bar (File, Edit, View, Terminal, Tabs, Help) and a root directory path "root@d8c06acf8f66: /work". The terminal shows a sequence of commands and their outputs: "ls" returns "rivet-plots", and "firefox rivet-plots/index.html" opens a browser window.

```
root@d8c06acf8f66: /work
andy@unity:~/tmp/docker$ ls
rivet-plots
andy@unity:~/tmp/docker$ firefox rivet-plots/index.html
andy@unity:~/tmp/docker$
```


Getting & using Rivet

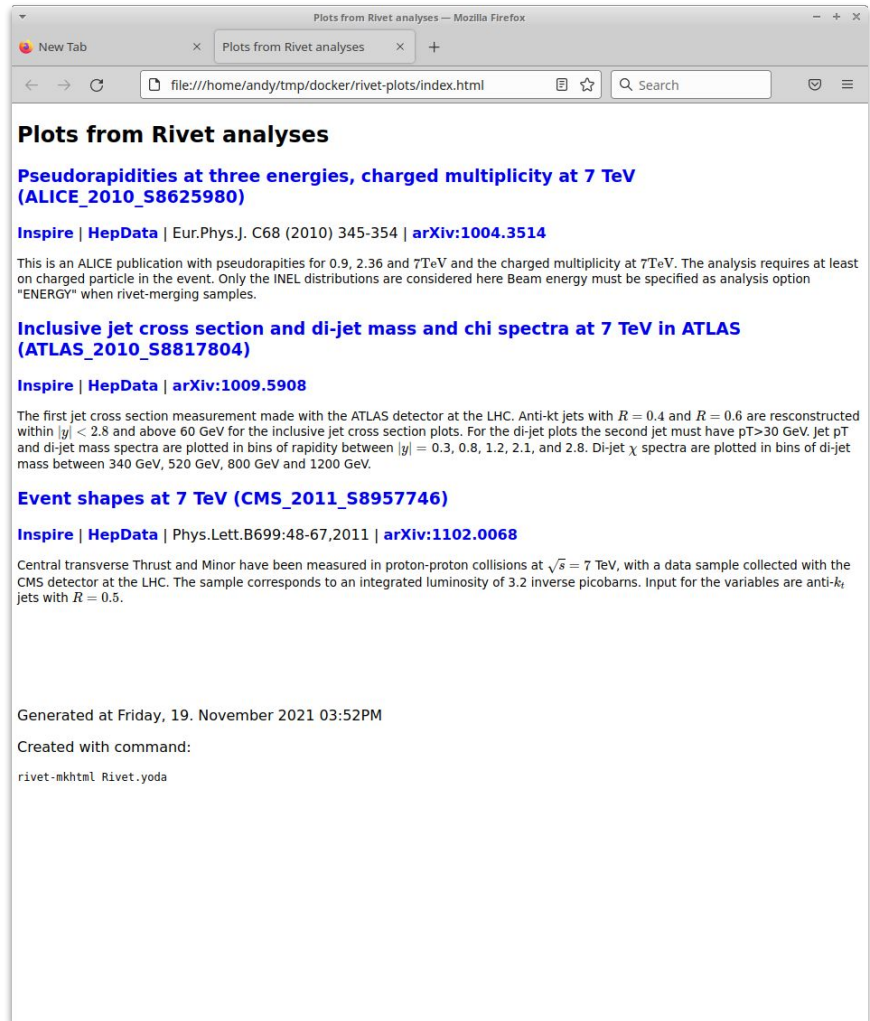
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery ⇒ copy an existing analysis!



The screenshot shows a Mozilla Firefox browser window with the address bar displaying `file:///home/andy/tmp/docker/rivet-plots/index.html`. The page content is titled "Plots from Rivet analyses" and lists several physics analyses with their corresponding journal references and arXiv IDs. The analyses include pseudorapidity distributions at 7 TeV for ALICE, inclusive jet cross sections and di-jet mass spectra for ATLAS, and event shapes for CMS. Each entry includes links to "Inspire" and "HepData".

Plots from Rivet analyses

Pseudorapidities at three energies, charged multiplicity at 7 TeV (ALICE_2010_S8625980)
[Inspire](#) | [HepData](#) | Eur.Phys.J. C68 (2010) 345-354 | [arXiv:1004.3514](#)

This is an ALICE publication with pseudorapities for 0.9, 2.36 and 7TeV and the charged multiplicity at 7TeV. The analysis requires at least one charged particle in the event. Only the INEL distributions are considered here Beam energy must be specified as analysis option "ENERGY" when rivet-merging samples.

Inclusive jet cross section and di-jet mass and chi spectra at 7 TeV in ATLAS (ATLAS_2010_S8817804)
[Inspire](#) | [HepData](#) | [arXiv:1009.5908](#)

The first jet cross section measurement made with the ATLAS detector at the LHC. Anti-kt jets with $R = 0.4$ and $R = 0.6$ are reconstructed within $|y| < 2.8$ and above 60 GeV for the inclusive jet cross section plots. For the di-jet plots the second jet must have $pT > 30$ GeV. jet pT and di-jet mass spectra are plotted in bins of rapidity between $|y| = 0.3, 0.8, 1.2, 2.1, \text{ and } 2.8$. Di-jet χ spectra are plotted in bins of di-jet mass between 340 GeV, 520 GeV, 800 GeV and 1200 GeV.

Event shapes at 7 TeV (CMS_2011_S8957746)
[Inspire](#) | [HepData](#) | Phys.Lett.B699:48-67,2011 | [arXiv:1102.0068](#)

Central transverse Thrust and Minor have been measured in proton-proton collisions at $\sqrt{s} = 7$ TeV, with a data sample collected with the CMS detector at the LHC. The sample corresponds to an integrated luminosity of 3.2 inverse picobarns. Input for the variables are anti- k_r jets with $R = 0.5$.

Generated at Friday, 19. November 2021 03:52PM

Created with command:

```
rivet-mkhtml Rivet.yoda
```

Getting & using Rivet

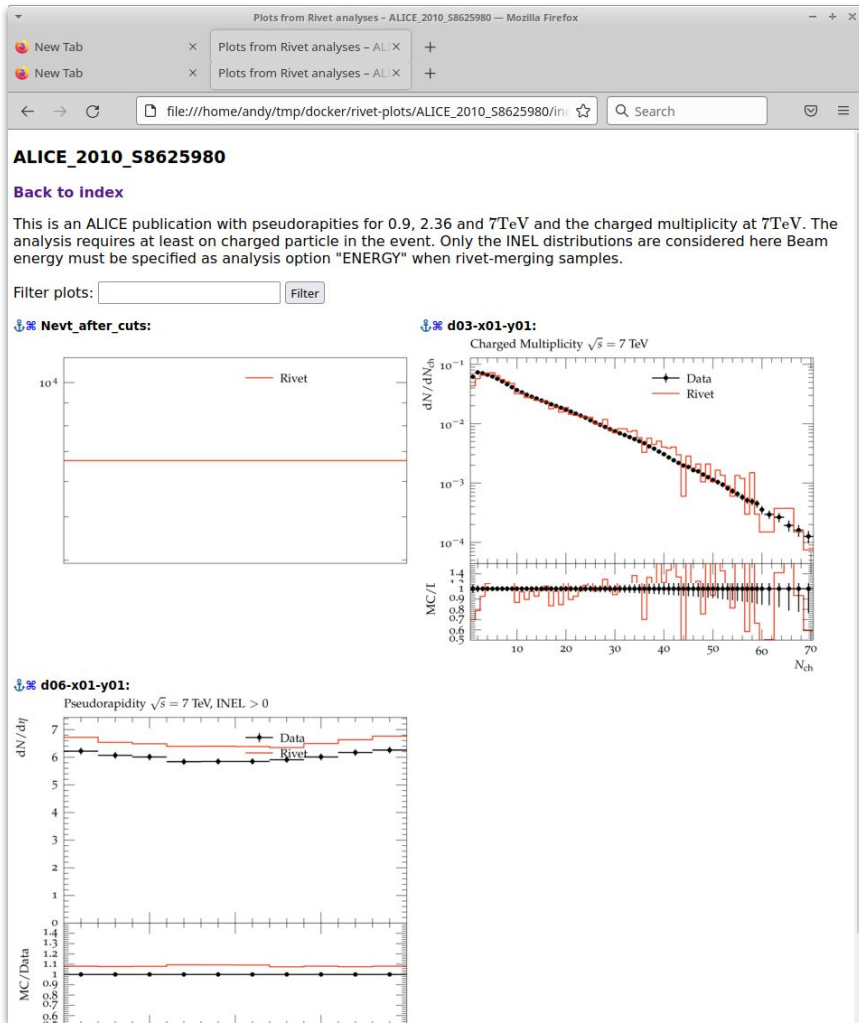
Lightweight analysis preservation
is valuable... and easy to start

As either a “user” or analysis
author, the barrier is lower than
ever: we recommend using our
Docker images to get started

Ideal for student projects!

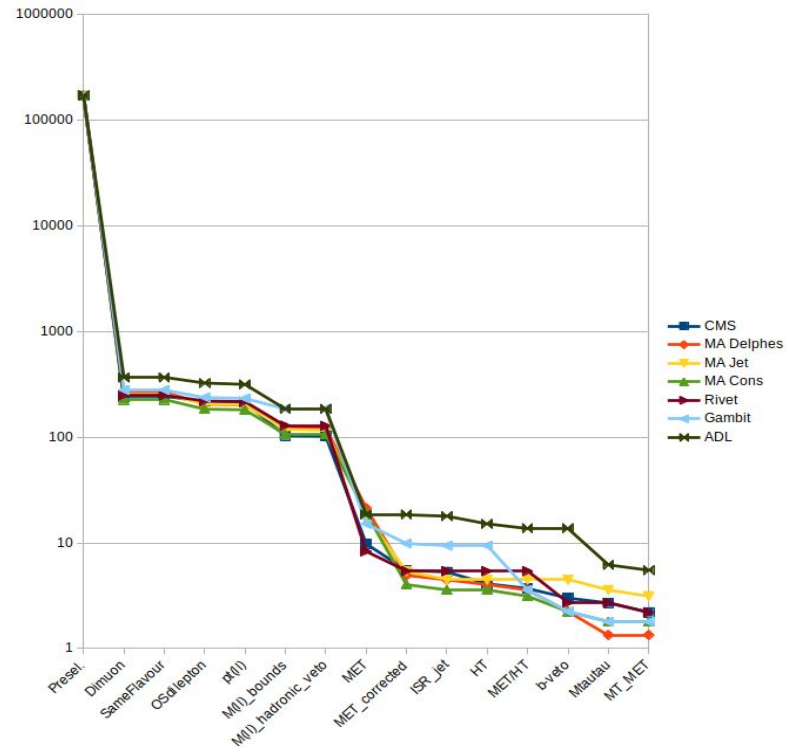
Tutorials available from the
[Rivet website](#), a **walkthrough** in
the [R3 paper](#)

Imitation the highest form of
flattery \Rightarrow copy an existing analysis!



Rivet and BSM-search recasting

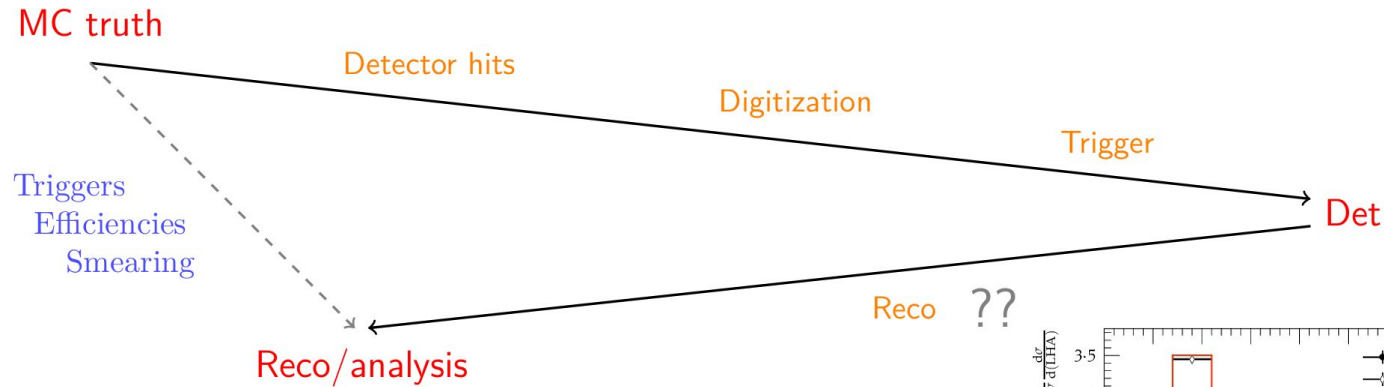
- ❖ Rivet's main emphasis *isn't* BSM direct searches, but there's no reason not to
 - lots of experiment experience and support
 - efficient scaling-up to hundreds of analyses, with distinct phase-space specific detector/efficiency functions
- ❖ Extra capabilities can lead to novel studies
 - new areas, collaborations, interested users...



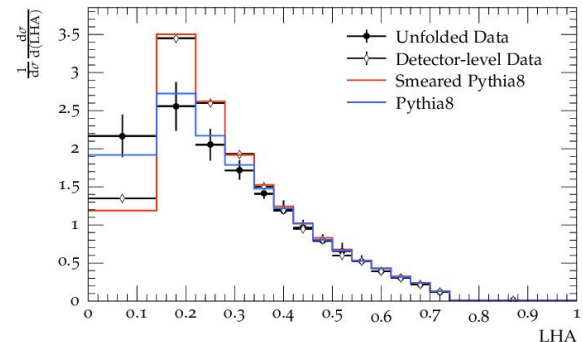
Detector emulation (but unfold by preference!)

❖ Detector smearing built on Rivet's projection system — for reco-level analyses

- developed based on GAMBIT ColliderBit experience: no need for “full fast-sim”



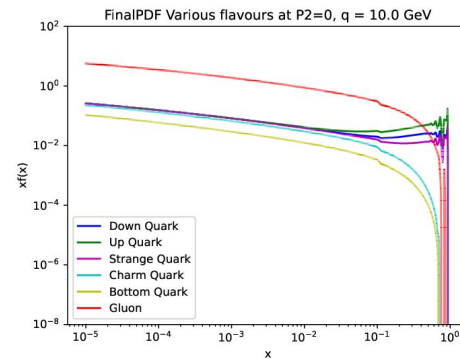
- like Delphes, but more flexible & can be *analysis-specific* ⇒ MA5 “SFS” mode
- flexibility allows e.g. “tuned” jet-substructure smearing, systematics studies, ...



The future of Rivet

- ❖ **Vision: Rivet as a standard for “truth-level” observables**
- ❖ Eyes on future colliders, including EIC, cosmic-ray air showers ... and nuclear physics, COMPASS? Happy to try!
- ❖ Not just standalone, but as a library in pheno & experiment frameworks, too: **leverage analysis collection, standardise MC-observable definitions, seamless systematics handling, etc.**
- ❖ Version 2 features include [high-dimensional \(and consistent\) histogramming](#), HDF5 aux data, and ONNX machine-learning.
- ❖ At its core: a **physics-oriented** system for physicists to **compare MC predictions to one another and to data, on many simultaneous observables, in myriad ways**

We don't know all the use-cases yet.



BSM from “Standard Model”

- ❖ **Not being focused on *direct* searches doesn’t mean no interest in BSM!**
- ❖ **Particle-level measurements *can* achieve high model-independence**
 - Careful definition of fiducial cross-section
 - Control distributions of “hidden variables” which are cut on
 - Reduce model sensitivity in unfolding
- ❖ **Rivet used directly in e.g.**
 - TopFitter top quark EFT fits;
 - at core of ATLAS *VH* EFT fits;
 - being integrated into Gambit global fits; and...
- ❖ **Contur is getting particular uptake**
 - Inject signal to “SM” measurements: **if it’d be statistically distinct, the model is eliminated**
 - Rivet gives huge coverage from “many angles”: views on not all, but most BSM signatures
a new result with Rivet code can be in Contur (or other) BSM fits within *hours*

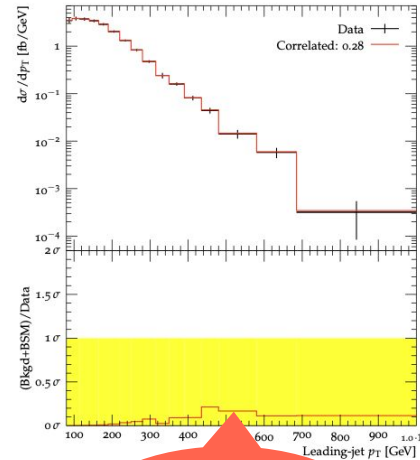
Contur



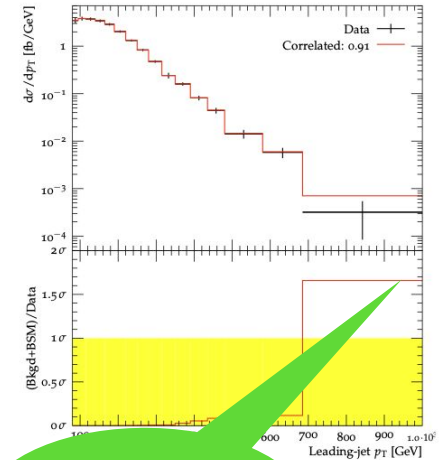
HT Louie Corpe

❖ Contur is “just” a wrapper on Rivet

- Ok, not just! You need to know which analyses are “safe”. Another reason for emphasis on final-states and *no cheating*
- In absence of unambiguous BSM, make zeroth-order assumption that data = SM
- Can be improved with high-precision SM theory predictions & uncertainties
- Signal-injection \Rightarrow care with e.g. ratios & profiles... cf. Rivet “perfect merging”
- Group analyses in stats-orthogonal “pools”. Use (expected) most-constraining element in the pool for setting limits — use correlations when possible to make “bigger” elements



Signal would have small effect wrt uncertainties, can't exclude it (28 % CL)



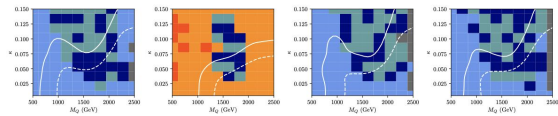
Signal would have large effects wrt uncertainties: can exclude at high CL

Try doing this with full-sim recast in finite time...

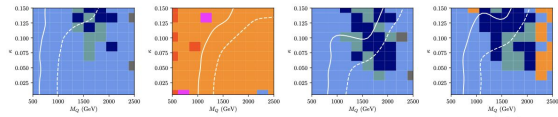
❖ Contur VLQ review requested a scan of realistic multiplets:

7 multiplets, each with 3 generational couplings, each with 4 W/H/Z-couplings,
300 points per scan, x 30,000 events \Rightarrow 750M events!

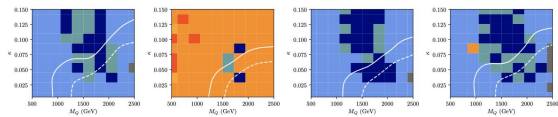
(a) BT 0:0:1 (b) BT 0:1:0 (c) BT 1:0:0 (d) BT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



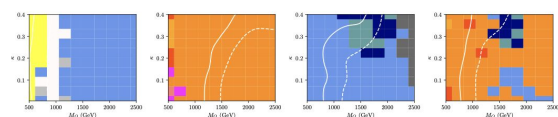
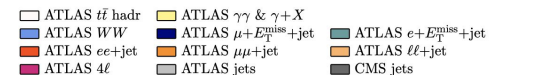
(e) XT 0:0:1 (f) XT 0:1:0 (g) XT 1:0:0 (h) XT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



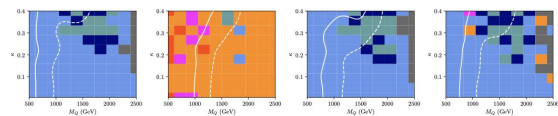
(i) BY 0:0:1 (j) BY 0:1:0 (k) BY 1:0:0 (l) BY $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



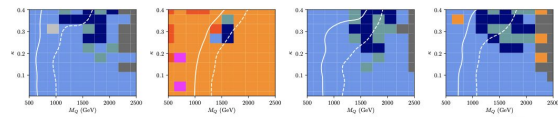
(m) $BTXY$ 0:0:1 (n) $BTXY$ 0:1:0 (o) $BTXY$ 1:0:0 (p) $BTXY$ $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



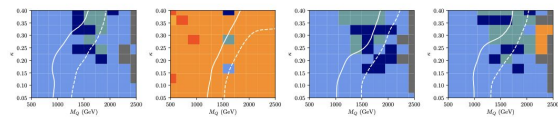
(a) BT 0:0:1 (b) BT 0:1:0 (c) BT 1:0:0 (d) BT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



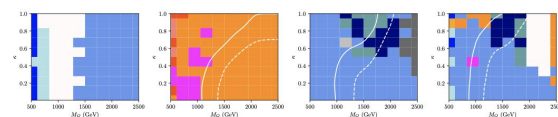
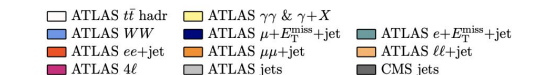
(e) XT 0:0:1 (f) XT 0:1:0 (g) XT 1:0:0 (h) XT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



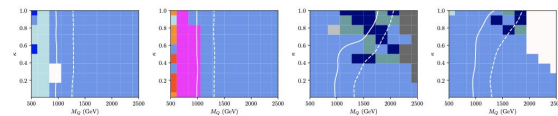
(i) BY 0:0:1 (j) BY 0:1:0 (k) BY 1:0:0 (l) BY $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



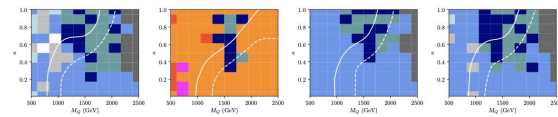
(m) $BTXY$ 0:0:1 (n) $BTXY$ 0:1:0 (o) $BTXY$ 1:0:0 (p) $BTXY$ $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



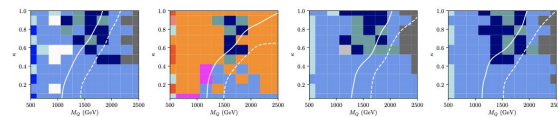
(a) BT 0:0:1 (b) BT 0:1:0 (c) BT 1:0:0 (d) BT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



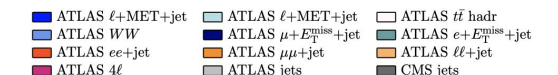
(e) XT 0:0:1 (f) XT 0:1:0 (g) XT 1:0:0 (h) XT $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



(i) BY 0:0:1 (j) BY 0:1:0 (k) BY 1:0:0 (l) BY $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



(m) $BTXY$ 0:0:1 (n) $BTXY$ 0:1:0 (o) $BTXY$ 1:0:0 (p) $BTXY$ $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$



Open questions / discussions on preservation

❖ Technical

- What obstacles to routine preparation in data / code ecosystem?
- ML preservations \Rightarrow validation and regression checks even more crucial (cf. ATLAS-SUS-2019-04 experience)
- Stability and longevity: how can we ensure the code still does the right thing in 2040? 2050? Beyond? Curation or containers?

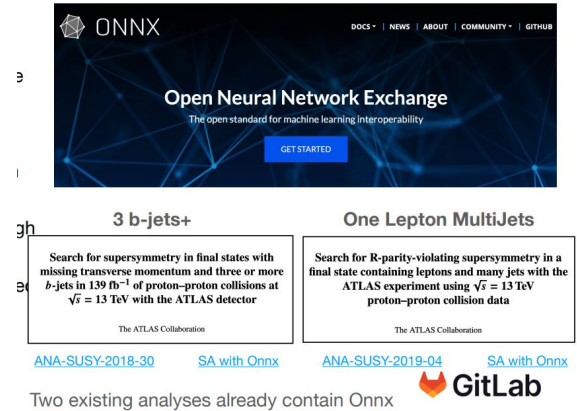
❖ Validation

- Huge pain / obstacle to preservation: figuring out the reference models, add cutflow checkpoints in the code (\rightarrow re-ordering...)
- Building cutflows / control distributions into analyses? Link to HepData? Ref HepMC samples on Zenodo?

❖ Social

- Valuing work that makes science reusable beyond The Paper
- Current (experimental) perceptions?
- Phenomenology input? Cross-community collaboration opportunities?

❖ AOB?



ONNX

DOCS • NEWS • ABOUT • COMMUNITY • GITHUB

Open Neural Network Exchange

The open standard for machine learning interoperability

GET STARTED

3 b-jets+

Search for supersymmetry in final states with missing transverse momentum and three or more b -jets in 139 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector

The ATLAS Collaboration

One Lepton MultiJets

Search for R-parity-violating supersymmetry in a final state containing leptons and many jets with the ATLAS experiment using $\sqrt{s} = 13 \text{ TeV}$ proton-proton collision data

The ATLAS Collaboration

[ANA-SUSY-2018-30](#) [SA with Onnx](#) [ANA-SUSY-2019-04](#) [SA with Onnx](#)

Two existing analyses already contain Onnx

