# MC analysis with Rivet

## Analysis prototyping, preservation & re-interpretation

**Andy Buckley,**
**University of Glasgow**

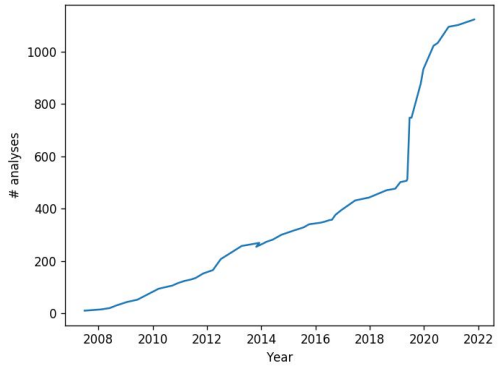


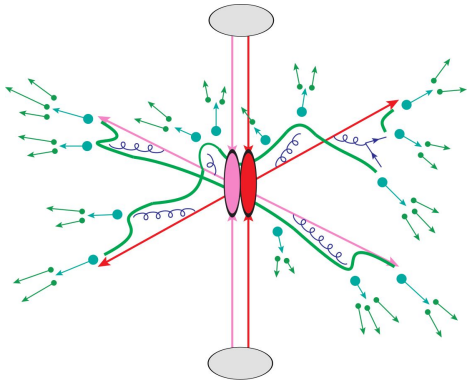
University of Glasgow

**23rd MCnet workshop**
**7 Dec 2021**

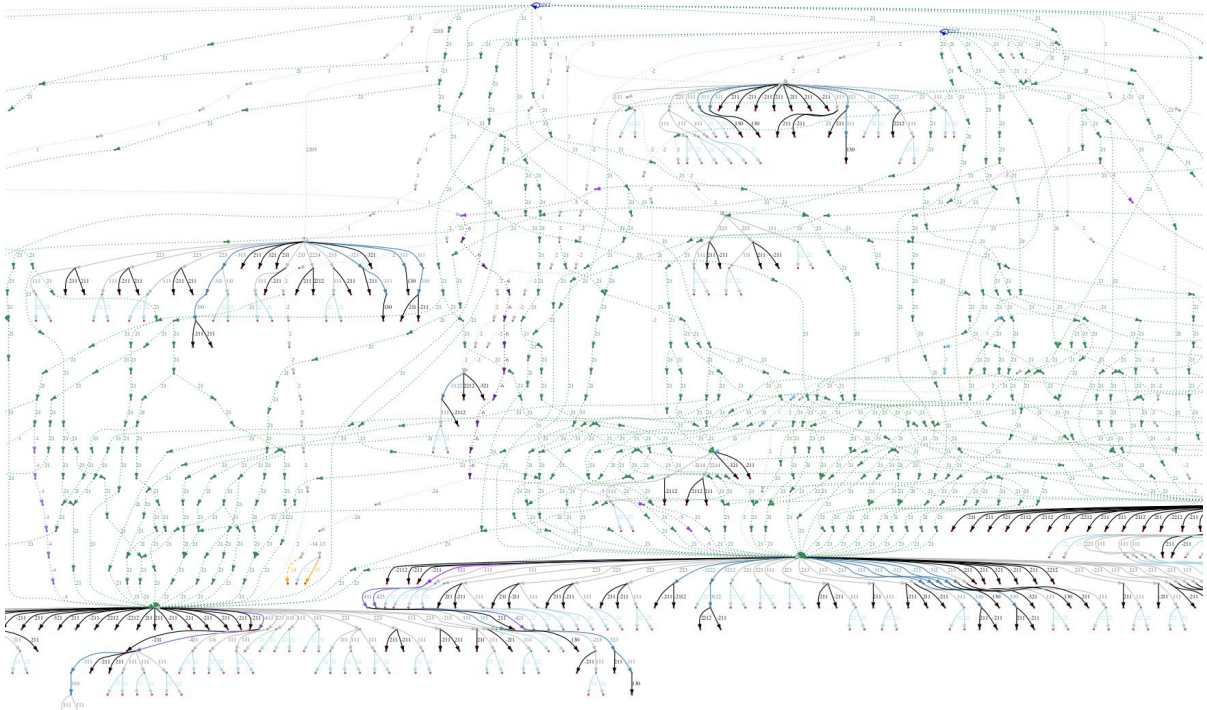
# What is Rivet?

- **The "LHC standard" MC analysis toolkit**
- **More broadly a project to preserve the logic of HEP data analyses and further expt-pheno collaboration**
  **⇒ the MCnet CEDAR meta-goal**
- **Code-wise, a C++ core and Python tools**
  - Fiducial / generator-independence emphasis
  - Integration with HepData
  - Transparent weight-stream handling
  - 1000+ analyses!
- **Central to a community of analysis reinterpretation tools, linking experiment to theory**
- **But why? Event loops are trivial…**



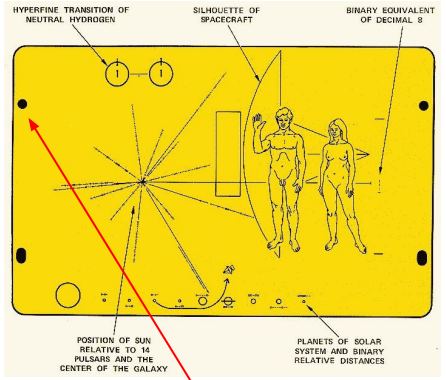



# Because of this:



We want to avoid all physicists needing to rediscover graph algorithms, conventions, pitfalls, physical/debug distinctions, …

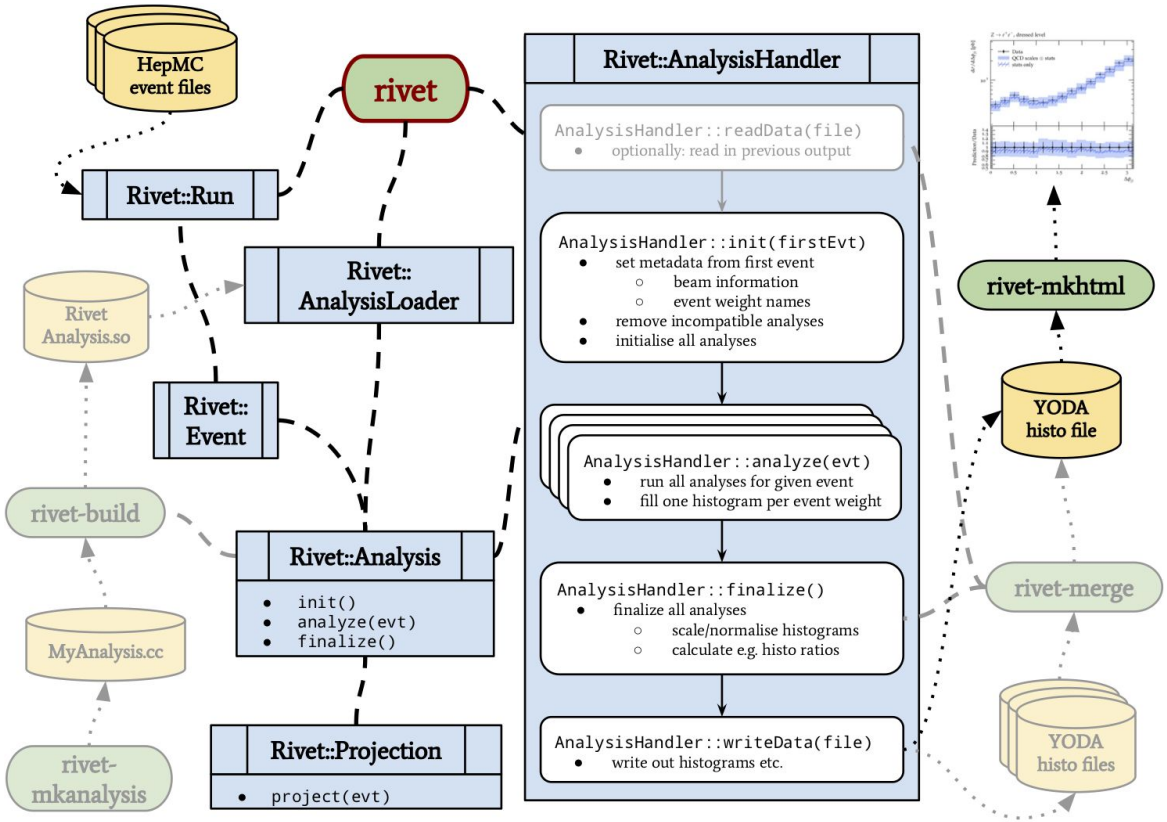
# Lessons learned




- **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!
    - (While I'm here… can't we make event records minimal by default?!)
  - 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

# The result

- **Rivet v3 structure**
  [arXiv:1912.05451](arXiv: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)

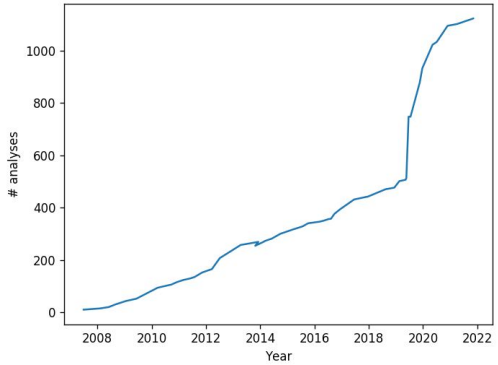


# The state we're in

- **Version 3.1.0 (2019) → 1000+ analyses!**

  A steady 50/yr flow of analysis submissions, plus a deluge of (mainly identified-hadron) routines from Herwig

  **⇒ v3.1.5 in Nov 2021**




- **Official support from the (LHC) experiments is crucial**

  preservation = standard part of "how we science", but still imperfect! We monitor coverage:

- **"New" features since the v1 vision:**

  systematics multiweights, "perfect merging", heavy ions, detector smearing functions, analysis options

**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 | $e^+e^-(\geq 12$ GeV) | $e^+e^-(\leq 12$ GeV) |
|---|---|---|---|---|---|---|---|---|
| **Rivet wanted (total):** | 72 | 111 | 126 | 183 | 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 greylist | Show blacklist

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

**ATLAS: Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment**
Inspire ID: 1802524 arXiv ID: 2006.13076 Report IDs: 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 $\sqrt{s} = 13$ TeV using th**
Inspire ID: 1801434 arXiv ID: 2006.09274 Report IDs: CERN-EP-2020-063
Links: Inspire CDS arXiv

**ATLAS: Measurements of the Higgs boson inclusive and differential fiducial cross sections in the $4\ell$ decay channel at $\sqrt{s}$ = 13 TeV**
Inspire ID: 1790439 arXiv ID: 2004.03969 Report IDs: CERN-EP-2020-035
Links: Inspire CDS arXiv HepData ATLAS_2020_I1790439

**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 IDs: CERN-EP-2020-030
Links: Inspire DOI/journal CDS arXiv HepData ATLAS_2020_I1790256

**ATLAS: Measurements of the production cross-section for a $Z$ boson in association with $b$-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS**
Inspire ID: 1788444 arXiv ID: 2003.11960 Report IDs: CERN-EP-2020-022
Links: Inspire CDS arXiv

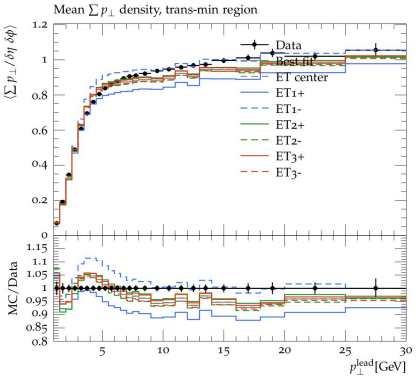
**ATLAS: Measurement of isolated-photon plus two-jet production in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector**
Inspire ID: 1772071 arXiv ID: 1912.09866 Report IDs: CERN-EP-2019-210
Links: Inspire CDS arXiv

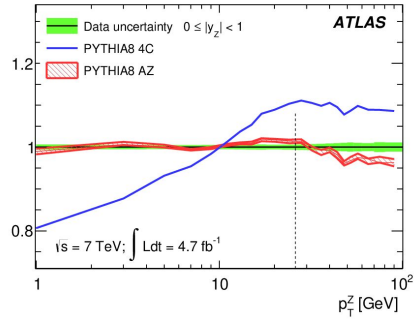
**ATLAS: A measurement of soft-drop jet observables in $pp$ collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV**

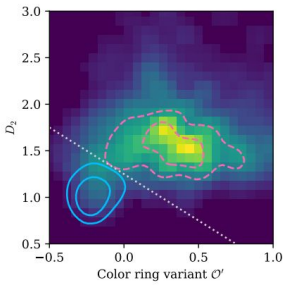
# Applications: from tuning to…

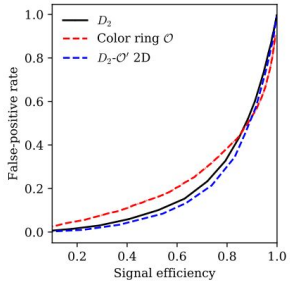
***First* "killer app": huge pre-LHC soft-QCD uncertainties:**

- **Tuning required Rivet analyses from expt**
- **Feed in to underlying event, pile-up, etc. modelling**
  - Better tunes ⇒ better analysis, better results
  - Impact: LEP and Tevatron analyses published for ~10 years suddenly got used! And cited…
    - ⇒ ATLAS tunes, CMS tunes, eigentunes
    - ⇒ Rapid responses to preliminary data
  - Model development: matching & merging, addition of energy evolution & colour-reconnection to Herwig, …
- Recently, also use of Rivet's large analysis collection for *BSM (see Contur) & Higgs*
  - Uptake still growing, e.g. in CMS



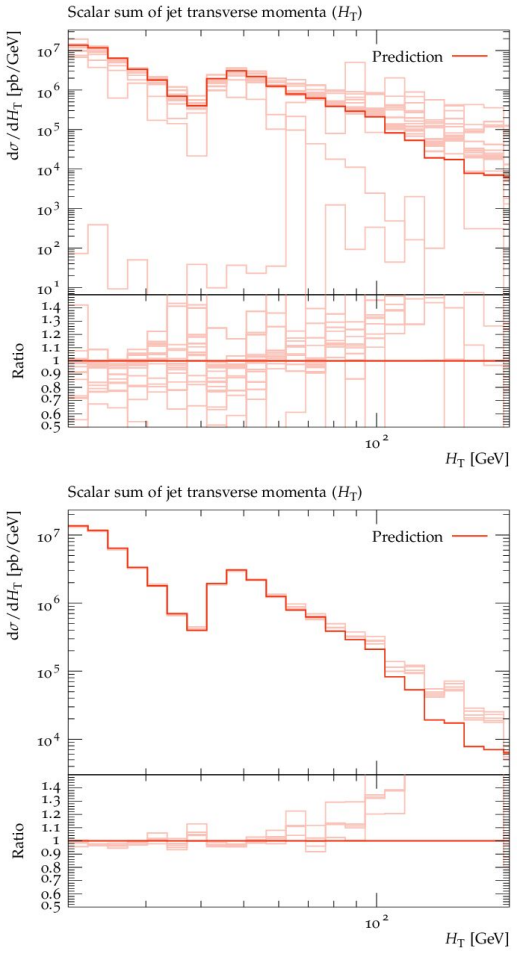



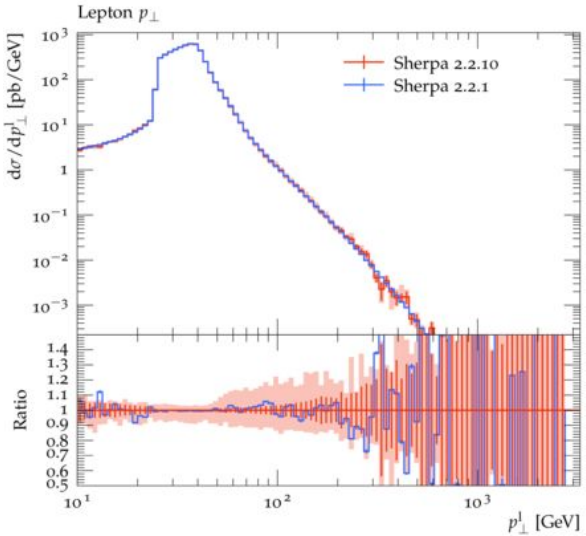
# Multiweights and re-entry

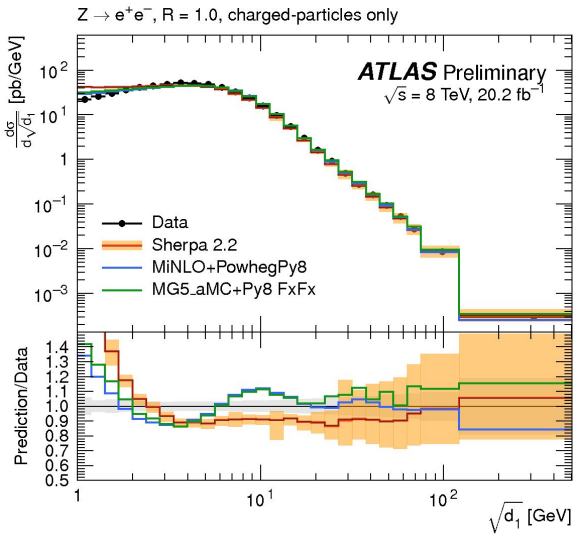
- **MC weight vectors allow expression of increasingly complex theory uncertainties.** But a burden for analysis chains: have to propagate and correctly combine O(200) weight streams!
- **Rivet 3: complex automatic handling of weights**
  ~invisible to users: data objects *look* like histograms etc. 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

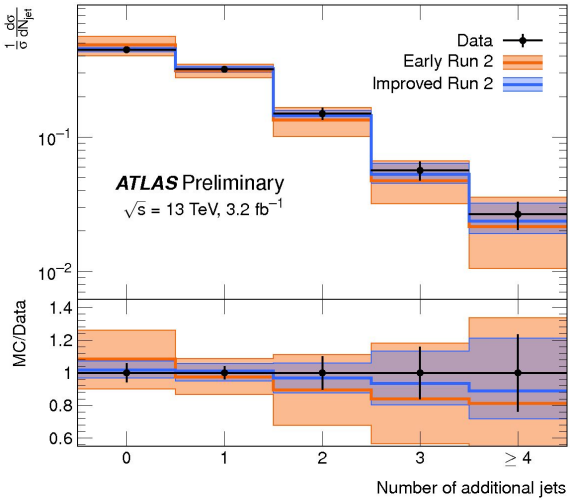


# MC systematics bands via multiweights

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



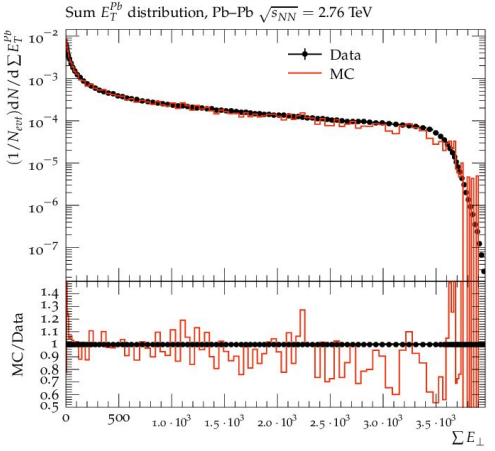



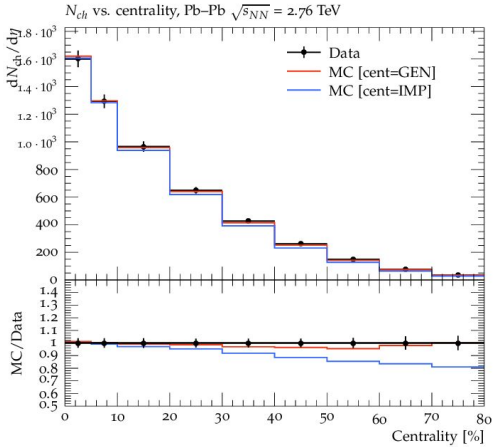




- Weight-naming standardisation ~~underway~~ concluding!

# Heavy-ion physics preservation

- "Adding heavy-ion support" sounds trivial!
- Actually nuanced ⇒ 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
- HI experience ⇒ updated *pp* primary particle defns



# HI community engagement!

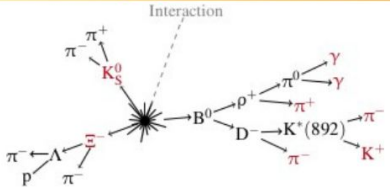
- **Great "spontaneous" engagement from within HI. Several productive workshops**

## Summary

1. Data getting into HEPData

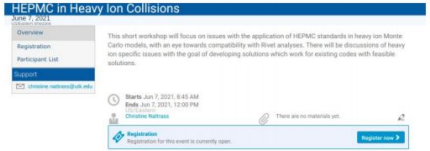
2. Build your own undergraduate army

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!
- https://indico.cern.ch/event/1022351/

# Also for EIC: *ep*/DIS/photoproduction

- **Recent pushes to include more *ep* analyses**
  - Remember Rivet's origins in HERA HZTool?
  - 2 analyses and DIS boosted frames since the beginning
  - Older attempt to port HZTool analyses to Rivet flopped: ~little interest → semi-useful RivetHZTool package
  - Now changing, largely due to EIC
- **8+2 new HERA routines in v3.1.5**
  - Supplied via preservation effort by Andrii Verbytskyi
  - More to come…
- **18 more from DESY summer students**
  - 2019 Rivet/DESY preservation workshop: refinements in fiducial machinery for DIS
  - Rivet-based preservation programme in summer 2021
  - Coming soon… v3.1.6 in ~Feb 2022?

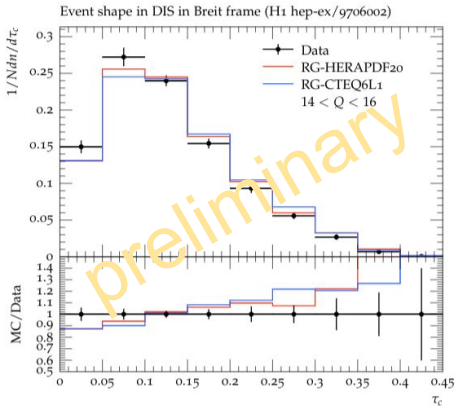
**Rivet, RivetHZTool and HERA**

**A validation effort for coding HERA measurements for Rivet**

Muhammad Ibrahim Abdulhamid[8], Andrea Achilleos, Giorgia Bonomelli, Aryan Borkar[14], Madhav Chithirasreemadam, Maksim Davydov[16], Keila Moral Figueroa, Alejandro B. Galván, Susie Kim, Kritsanon Koenno[illegible], Luca Marsili[13], Ariadna León Quirós, Narmin Malimova, Jacob Shannon, Suraj Singh[14], Can Süslü[12], Natthaporn Trakulphorm, Danielle Wilson [illegible], Wenting Zhang

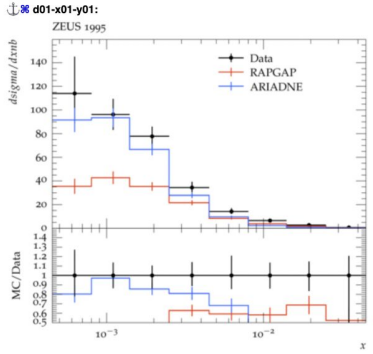
A. Bermudez Martinez[2], A. Buckley[9], C. Bierlich[15], J. M. Butterworth[3], L.I. Estevez Banos[2], C. Gütschow[11], H. Jung[2], M. Mendizabal[2], S. Taheri Monfared[2], S. Plätzer[10], S. Schmitt, P. van Mechelen[4], Q. Wang[2,7], G. Watt[17], M. Wing[11], H. Yang[2,7]

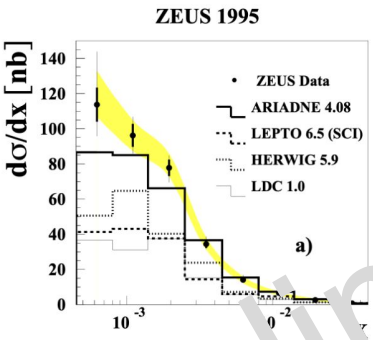


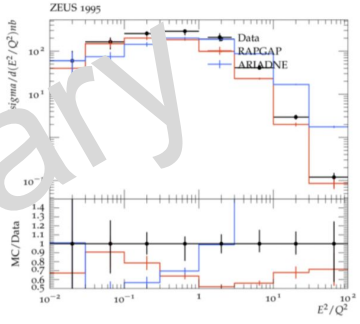


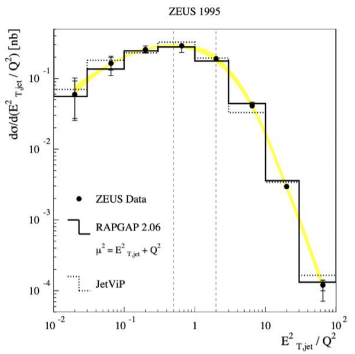
# More prelim HERA *ep* outputs

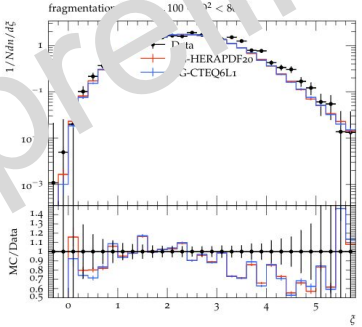
**Analyses on everything from jet rates, to event shapes, $E_T$ flow, *K* rates, *b* production, *D* and inclusive fragmentation, …**

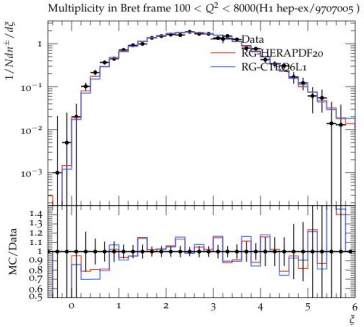


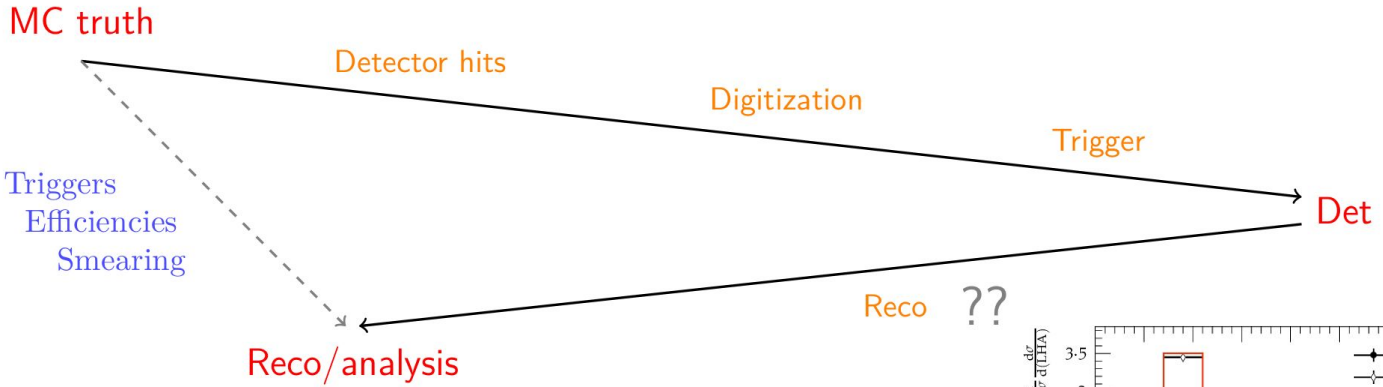



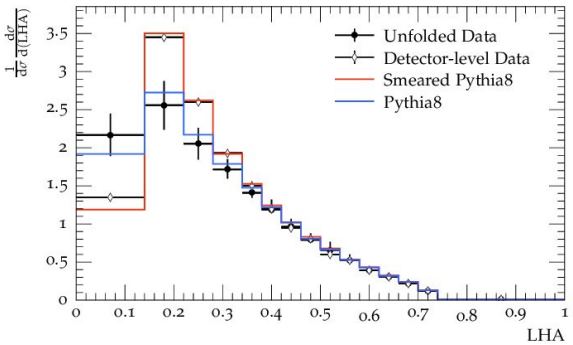



# Detector emulation (but please unfold by preference!)

- **Detector smearing built on Rivet's projection system — for reco-level analyses**
  - developed based on Gambit ColliderBit experience: no need for "full fast-sim"

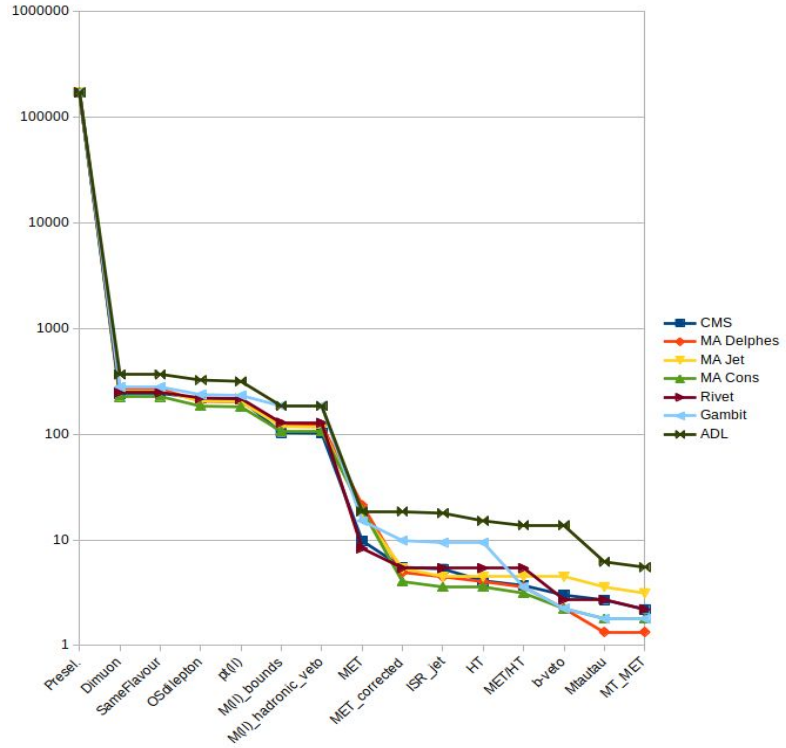



  - like Delphes, but more flexible & can be *analysis-specific* ⇒ MA5 "SFS" mode
  - flexibility allows e.g. "tuned" jet-substructure smearing, systematics studies, …



# 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
- **Can we do for BSM preservation what we did for measurement analyses?**
  - Hasn't been a major focus: mechanisms are useful anyway, experiments (at least ATLAS) have focused on home-grown solutions
  - But… recent CMS interest via Si Hyun Jeon. Watch this space 👍




Les Houches 2019 CMS soft-lepton recasting-tools comparison

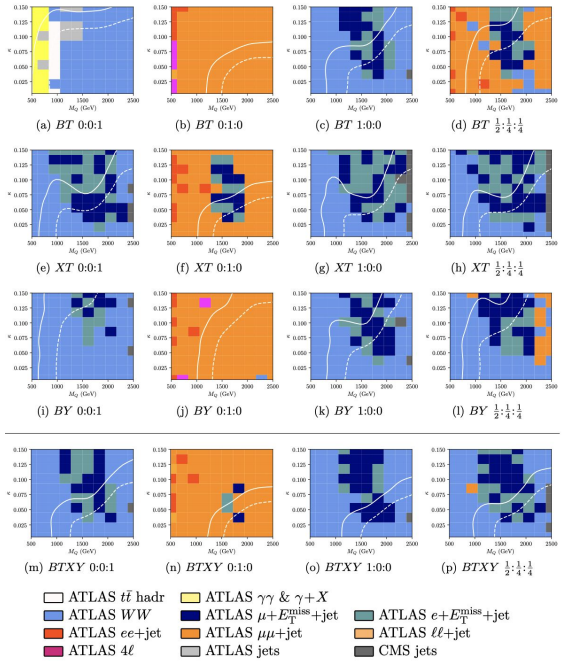
# BSM from “Standard Model”

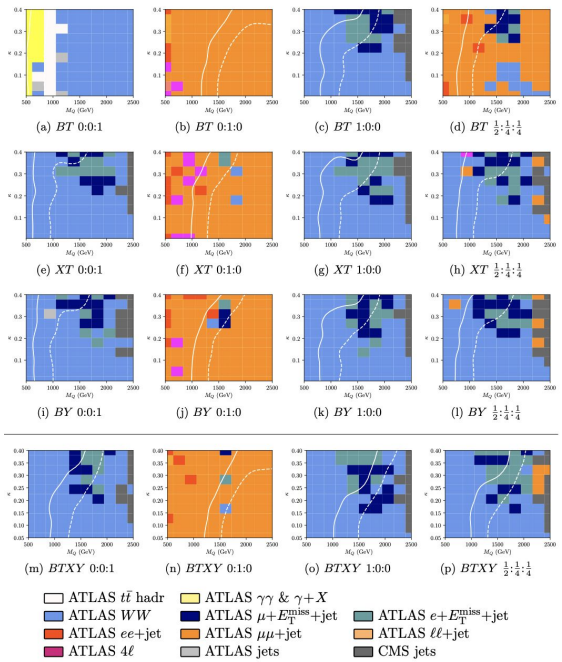
- **Not being focused on *direct* searches doesn’t mean no interest in BSM!**
- **Particle-level measurements *can* achieve high model-independence**
  - Careful definition of fiducial cross-section
  - Control distributions of “hidden variables” which are cut on
  - Reduce model sensitivity in unfolding
- **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

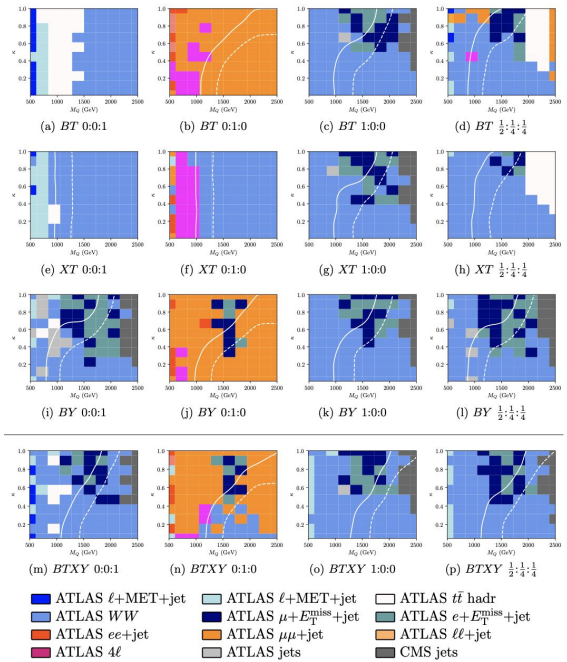
    a new result with Rivet code can be in Contur (or other) BSM fits within *hours*

# 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 ⇒ 750M events!



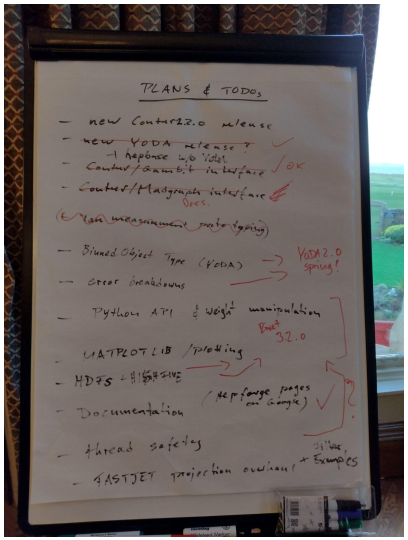
# The future of Rivet

- **Vision: Rivet as a standard for "truth-level" observables, across collider physics**
  - ➢ Already used this way inside CMSSW truth definitions 🤯
- Not just standalone, but as a library in pheno & experiment frameworks, too: leverage analysis collection, standardise MC-observable definitions, seamless systematics handling, etc.
- 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.

# Practical tasks and challenges

- **Tasks:**
  - Extension of HepData and other community infrastructure for ever-more precise data. Even our compressed data format is struggling with the volume of analyses and data
    GSoC+follow-up on generalised binned containers, static/dynamic object distinction, and multiweight-oriented data formats (HDF5)
  - Improved, modernised visualisation and exploration
    ⇒ matplotlib GSoC+follow-up to make public
  - Preserving MVAs: BDT and NN in vanilla C++? Or avoid?
- **Challenges:**
  - So much progress has happened at/because of in-person developer workshops ⇒ Covid had a big impact. Events in Dec 2020 and more recently have re-invigorated developments
  - Need to find ways to continue this without MCnet funding…




# Getting & using Rivet

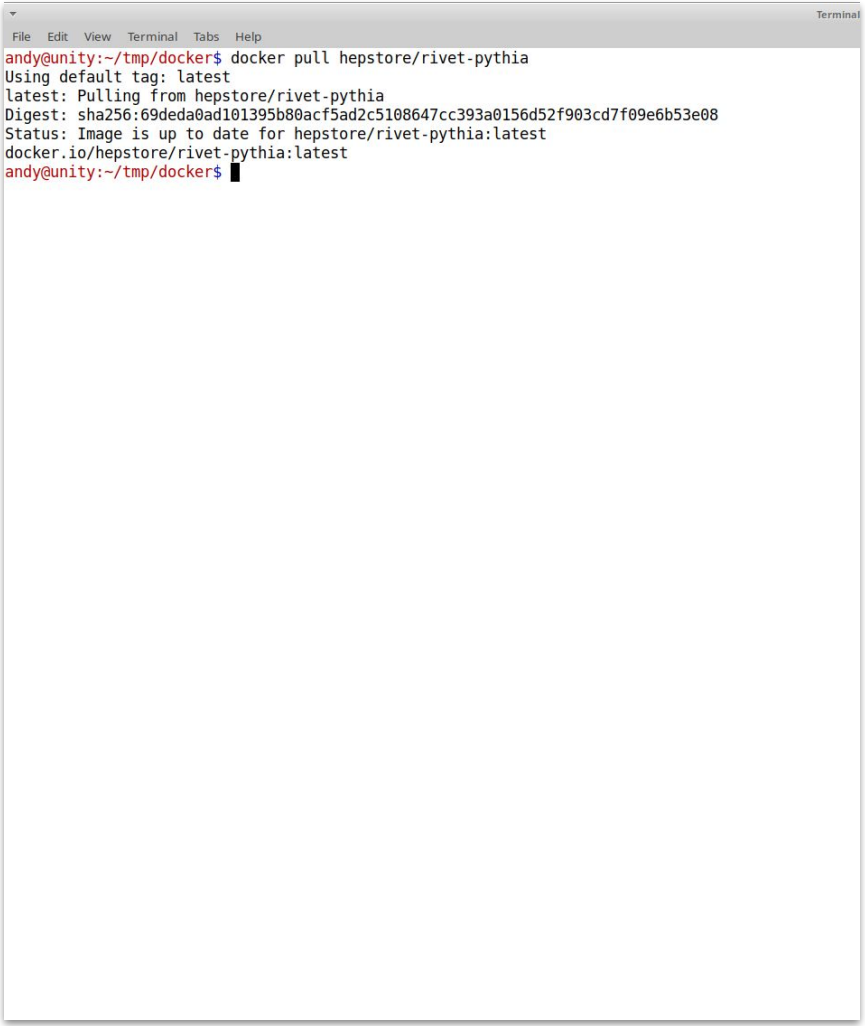
**Lightweight analysis preservation is valuable… and easy to start**

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!



# Getting & using Rivet

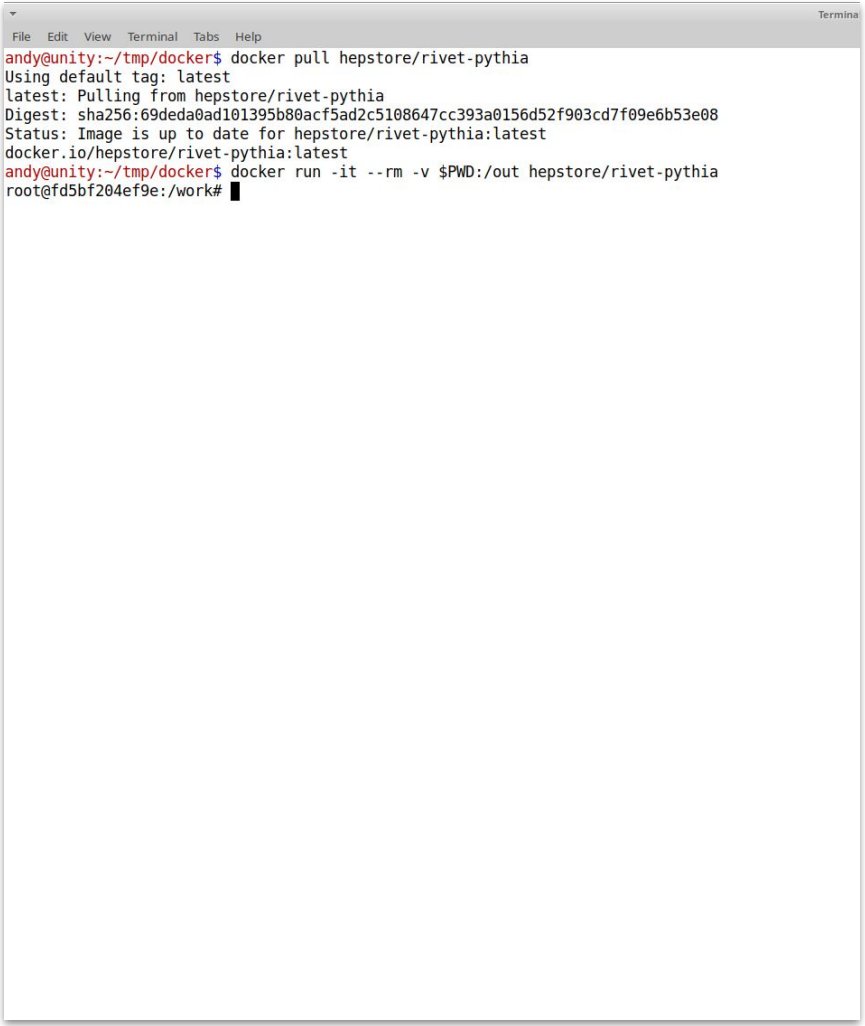
**Lightweight analysis preservation is valuable… and easy to start**

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!



# 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!

# Getting & using Rivet

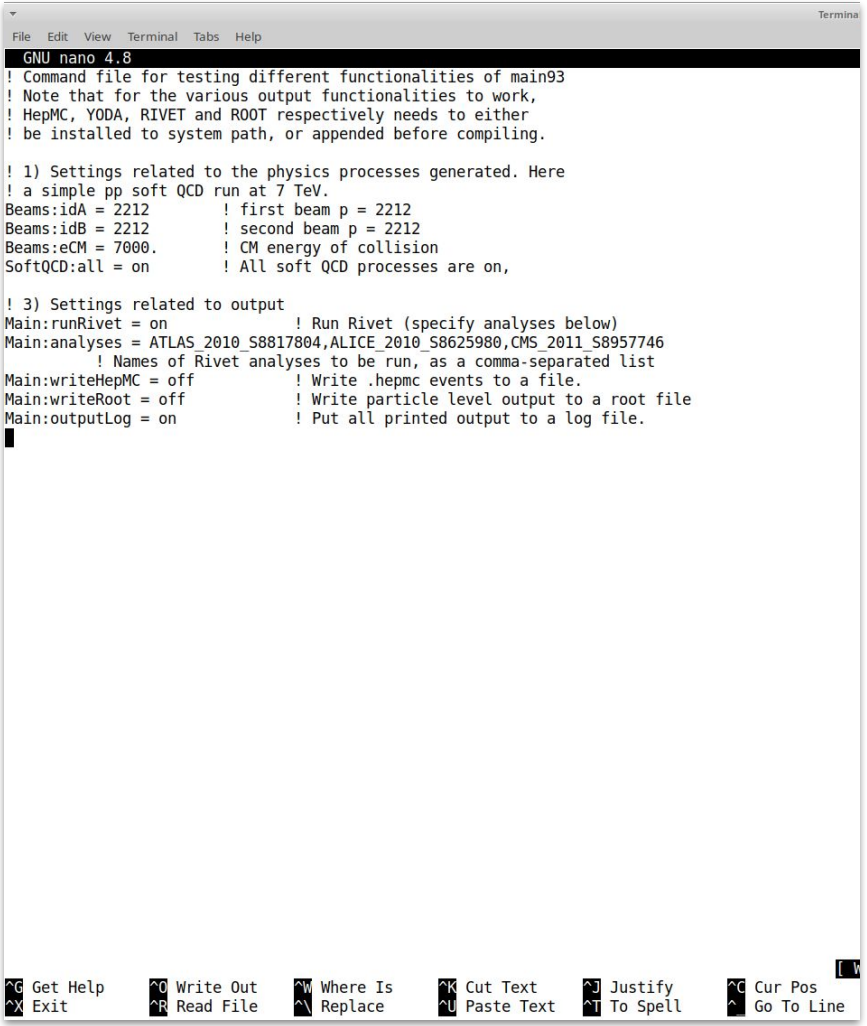
**Lightweight analysis preservation is valuable… and easy to start**

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!



# 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!

# 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!

# 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!

# 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!

# Getting & using Rivet

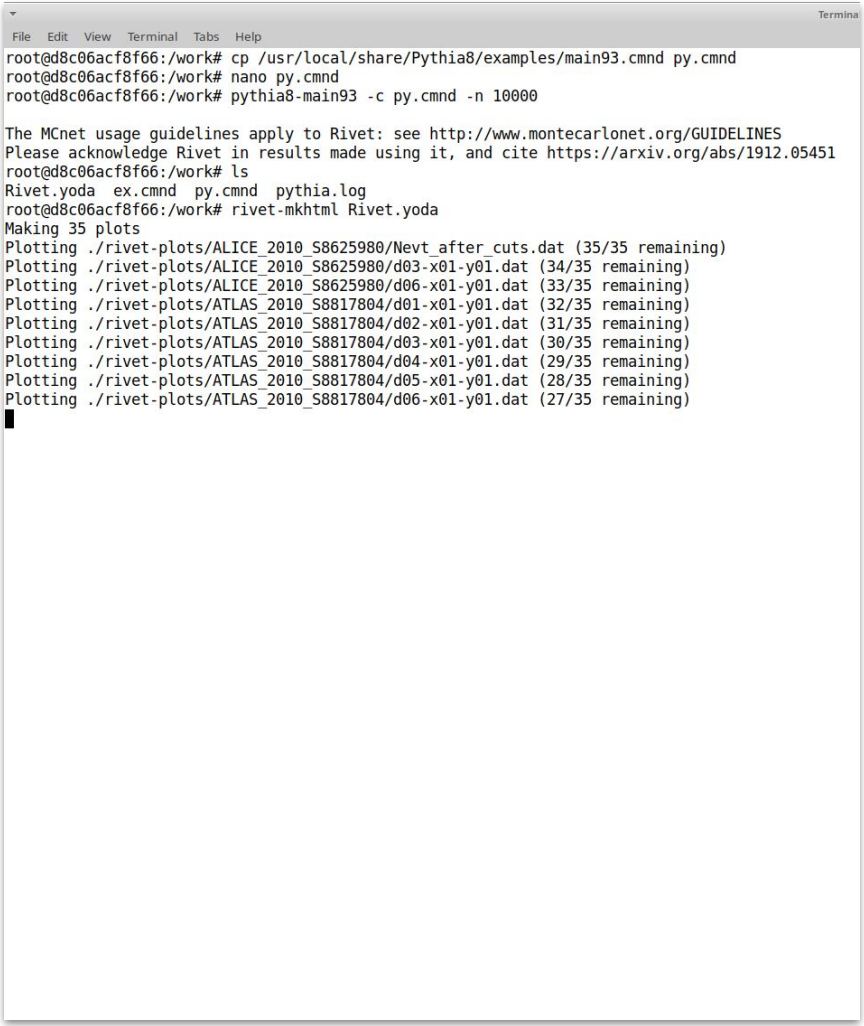
**Lightweight analysis preservation is valuable… and easy to start**

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!



# Getting & using Rivet

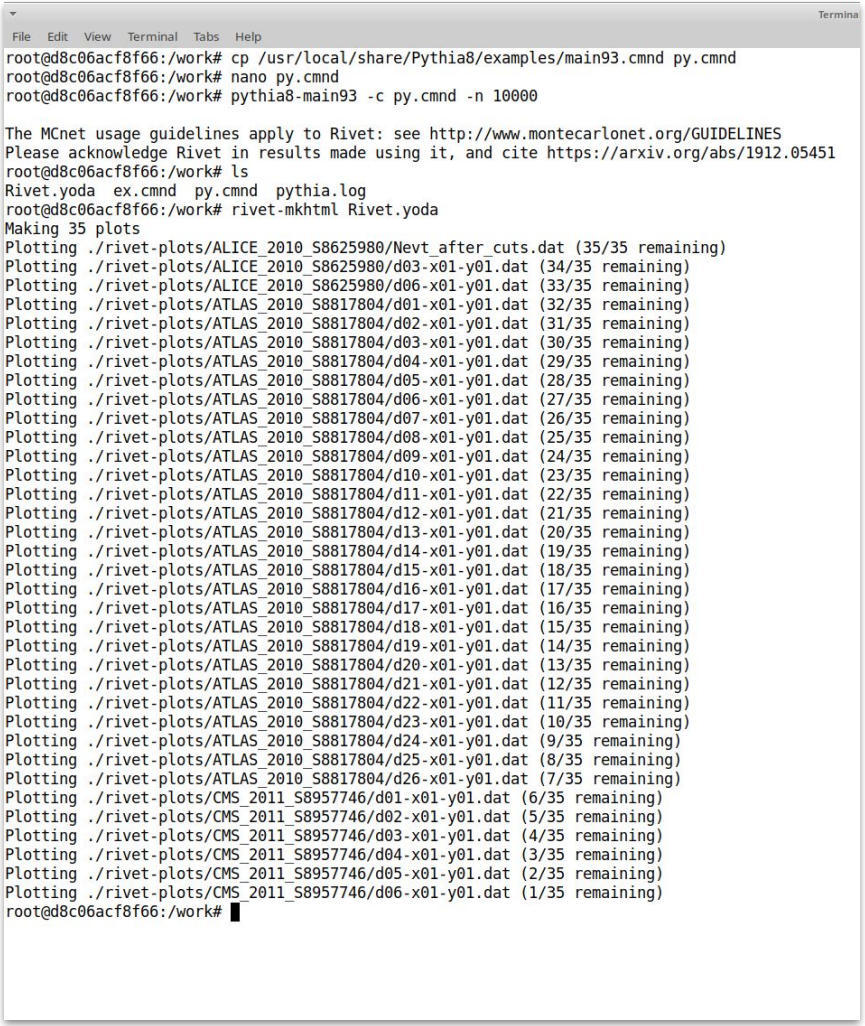
**Lightweight analysis preservation is valuable… and easy to start**

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!



# Getting & using Rivet

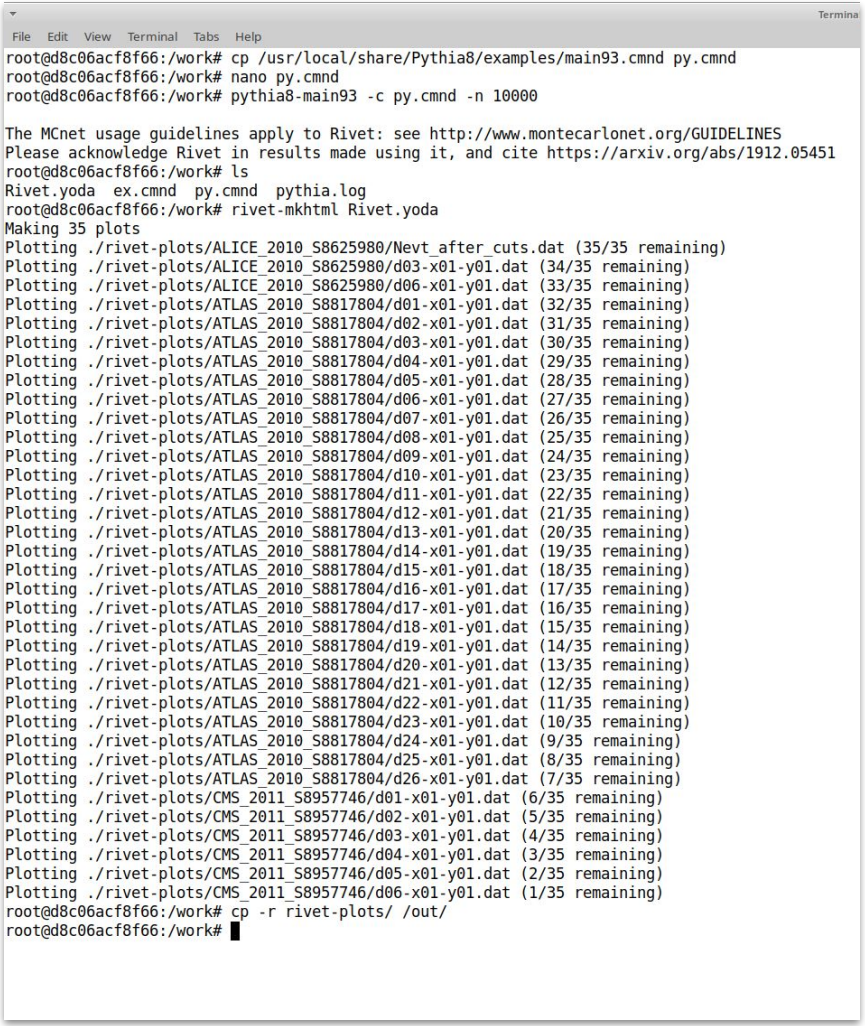
**Lightweight analysis preservation is valuable… and easy to start**

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!



# Getting & using Rivet

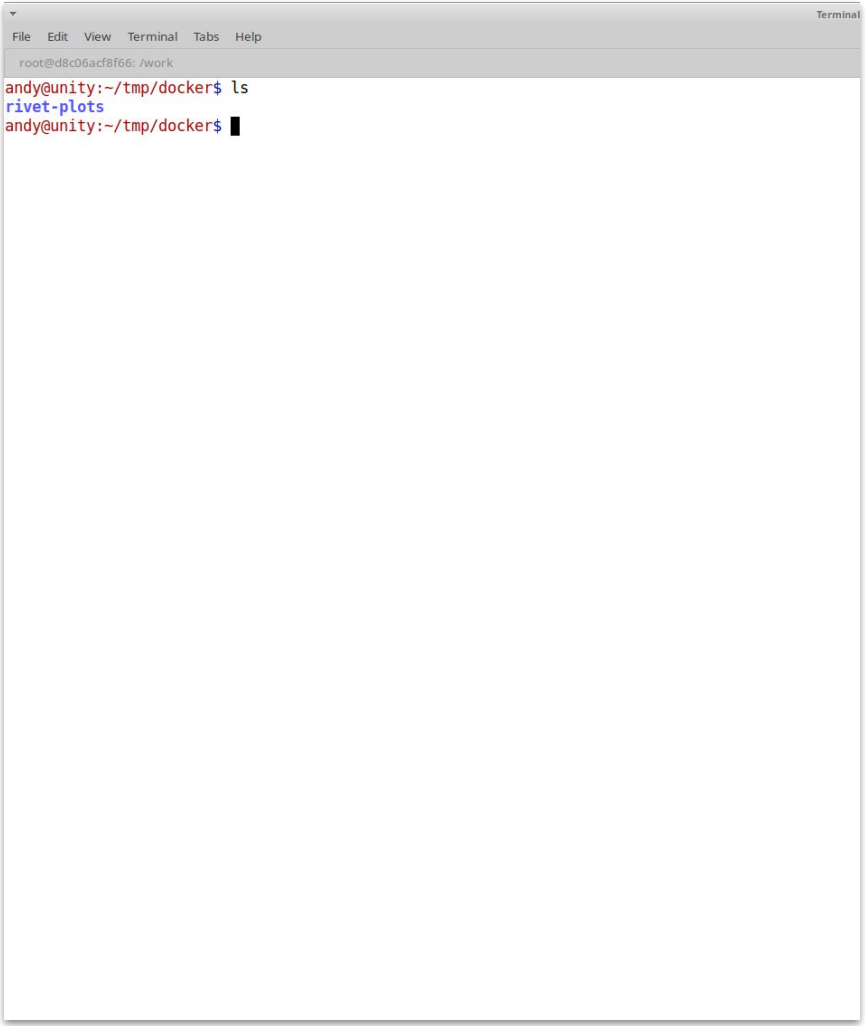
**Lightweight analysis preservation is valuable… and easy to start**

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!



# 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!

# Getting & using Rivet

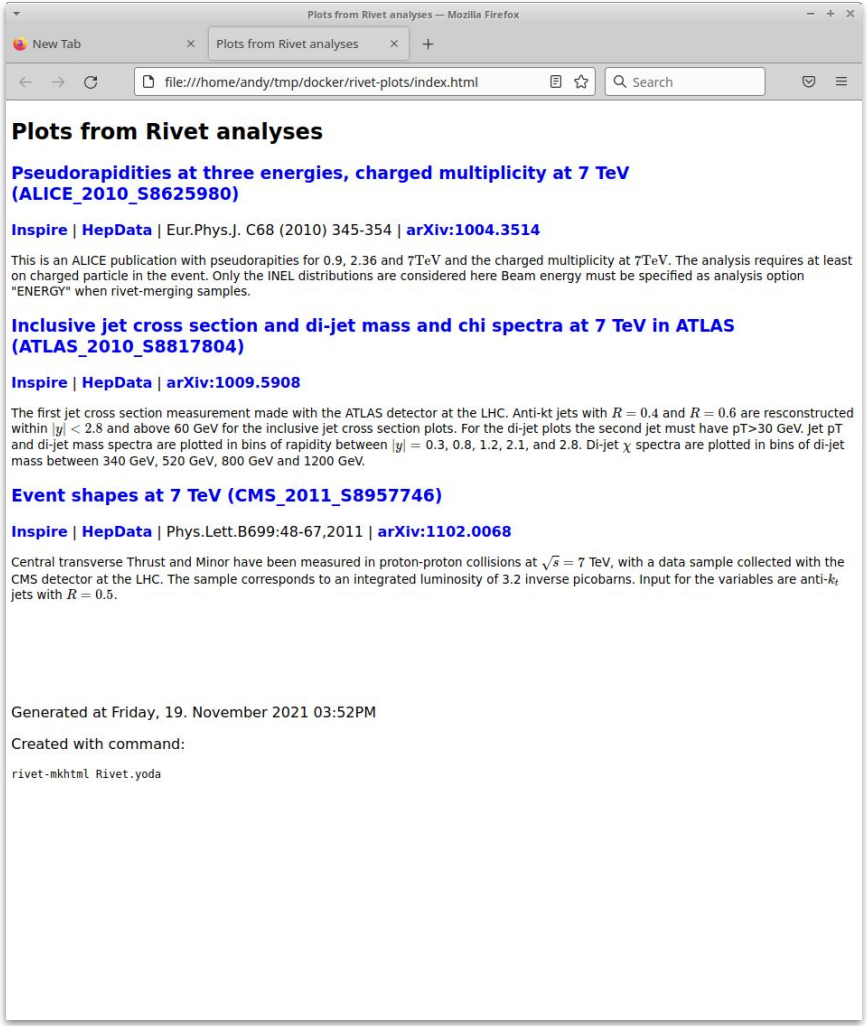
**Lightweight analysis preservation is valuable… and easy to start**

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!



# Getting & using Rivet

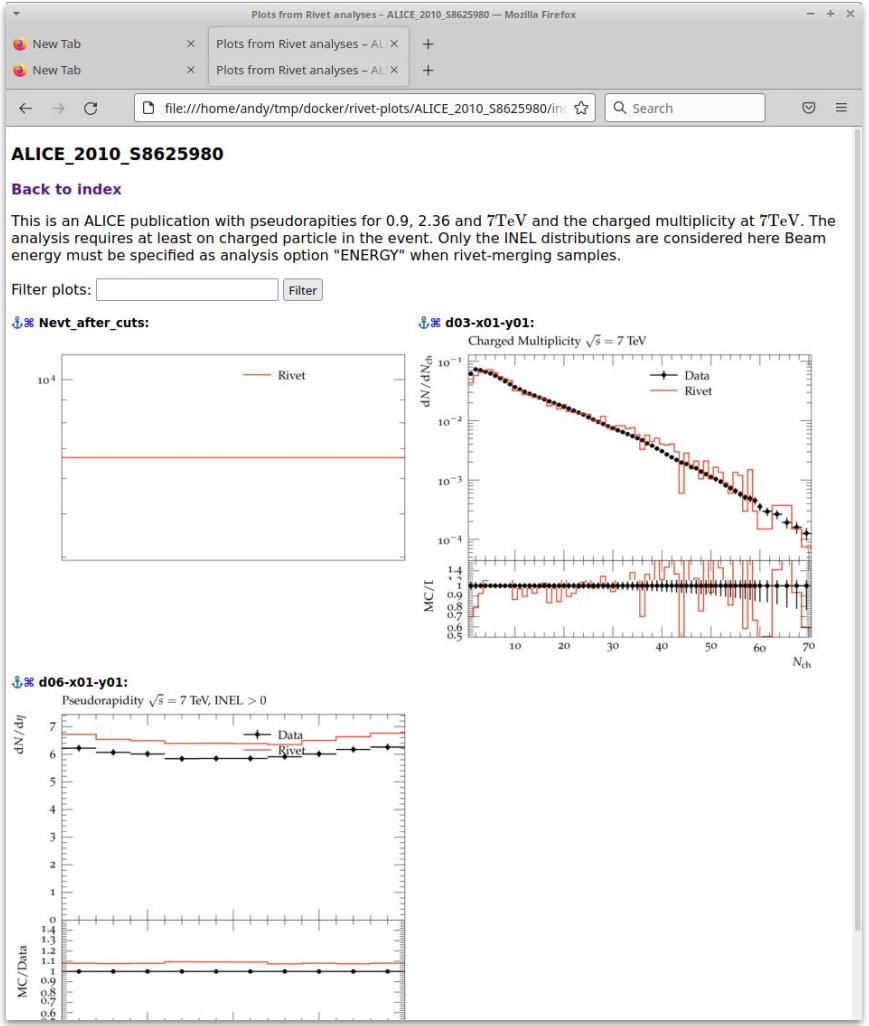
**Lightweight analysis preservation is valuable… and easy to start**

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!



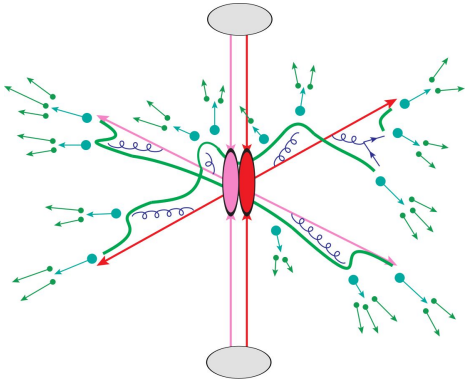
# Summary

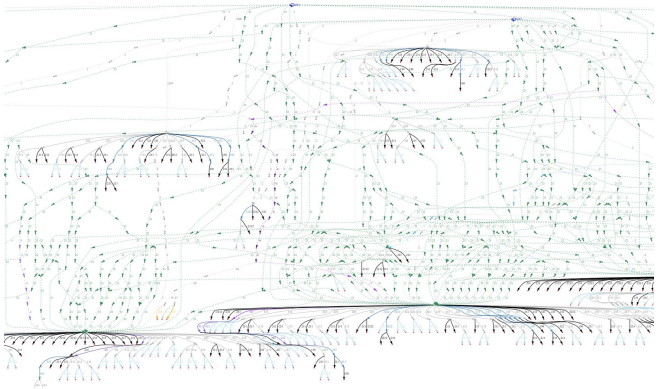
- **Rivet arose from HERA experiment/MC author collaboration, in time for the LHC**
- **Experiment/pheno activities resulted, e.g.**
  - ➢ MC(net) development, MC tuning, Higgs & BSM fits, …
- An accelerator for analysis impact: many exp/theory studies using Rivet as their common language
- Development during MCnet3 took us from v1 + ideas to next-gen tool. Collaboration ⇒ better ideas
- Now entering another new era: precision machinery for LHC is in place ⇒ need scaling & comprehensive uptake. And looking *beyond LHC*, to HI, EIC, astroparticle, …

# Backup slides

# MC generation

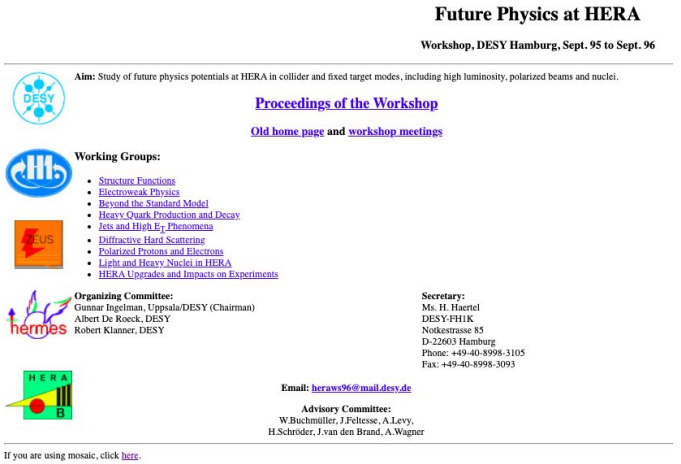
- **MC generation is where theory meets experiment**
  - The fundamental *pp, pA, AA* collision, *sans* detector
- **Components of an "exclusive" event-generator chain:**
  - QFT matrix element sampling at fixed-order in QCD
  - *Dressed* with approximate collinear splitting functions, iterated in factorised Markov-chain "parton showers"
  - FS parton evolution terminated at Q ~ 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





# 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 $Q^2$ ~ 4 GeV$^2$; jet $p_T$ ~ 5 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



**Future Physics at HERA**

**Workshop, DESY Hamburg, Sept. 95 to Sept. 96**

**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 $E_T$ 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:** heraws96@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.

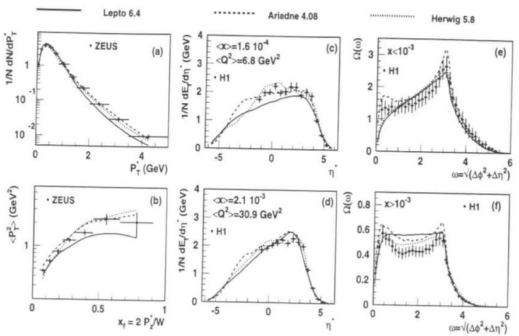
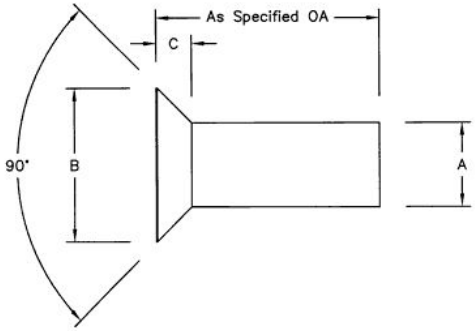



Figure 1: *The transverse momenta $dN/dp_T$ (a) and the 'seagull' plot $\langle P_T^2 \rangle \times x_F$ (b) of single particles in the positive hemisphere of the hadronic center of mass. The transverse energy flow $dE_T/d\eta$ in a low (c) and high (d) $x$ and $Q^2$ bin. The transverse energy-energy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

# Designing Rivet

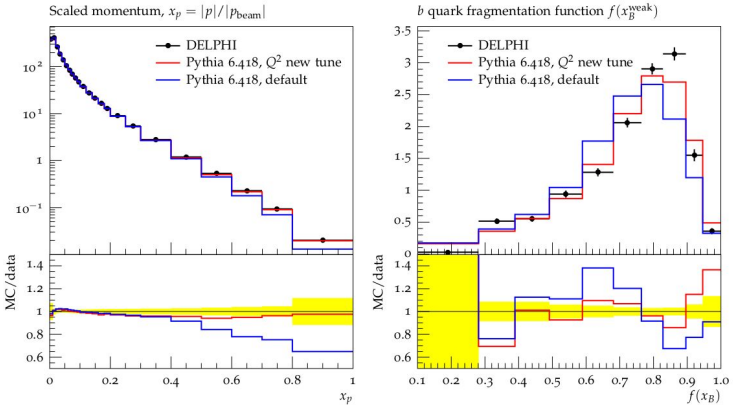
- **Ease of use**
  - **Big emphasis on “more physics, less noise”!**
  - Minimal boilerplate analysis code, HepData sync
  - Event loop and histogramming basically familiar
  - Tools to avoid having to touch the raw event graph
- **Embeddable**
  - OO C++ library, Python wrapper, sane user scripts
  - 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**

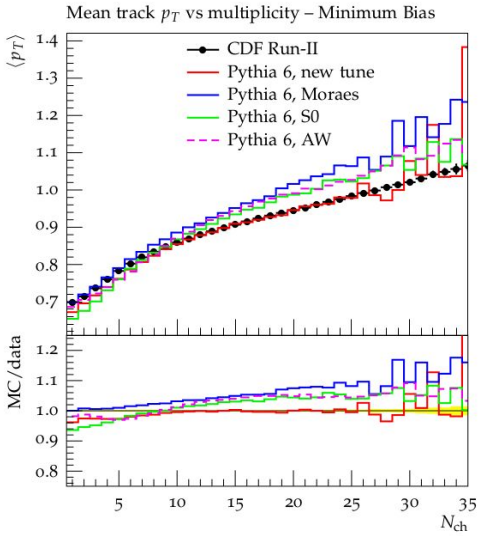



# Event-generator tuning

**Event generators all have dirty secrets. Usually non-perturbative ones… O(30+) parameters**

- **First systematic hadron collider "tunes" of PYTHIA6 by Rick Field for CDF ~ 2001**
  - Tune A, Tune D, Tune DW, etc. etc.
- **Limited datasets, variation by hand**
  - Rivet and its analyses were a game-changer
  - You only know a model is incapable when you've scanned its whole param space… and then the argument is over
- **The "Professor" tunes, 2008; and…**

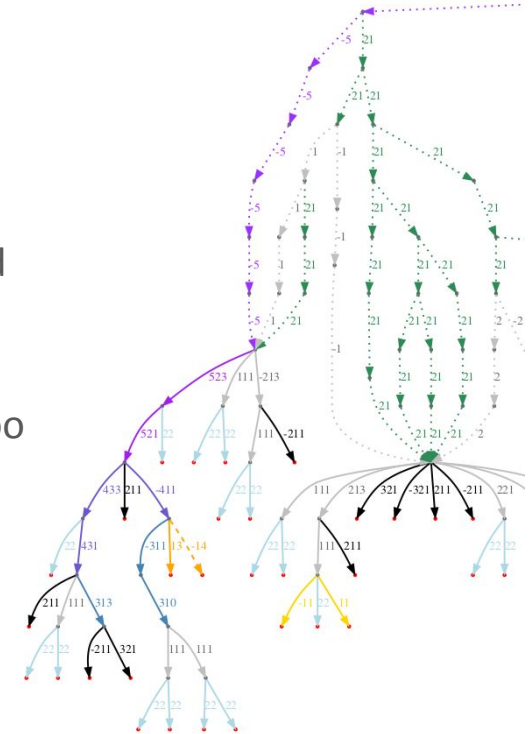


# Physically safe analysis methods


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

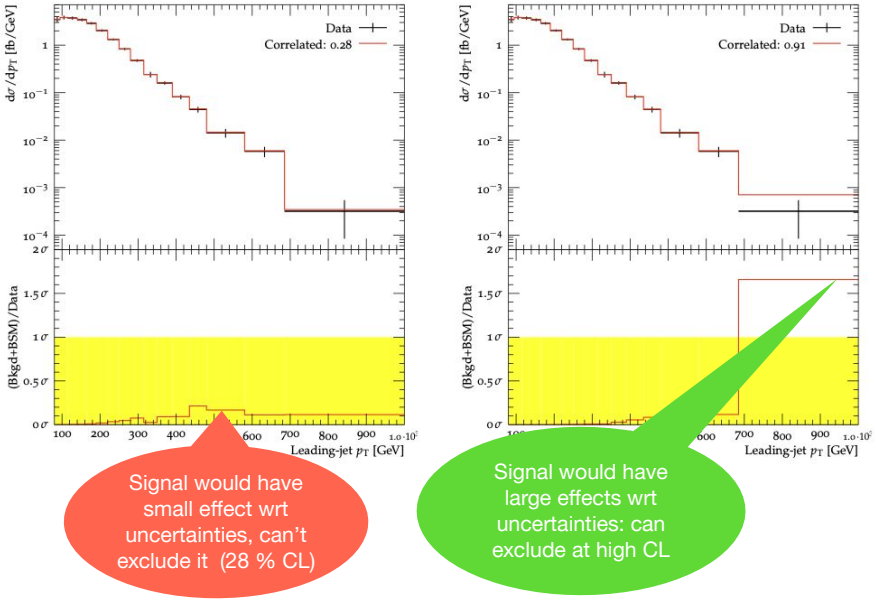
- **refining the "fiducial" idea, defining *unfolding targets***
- **Hadronisation as a "decoherence barrier"**
  use the natural dividing line between the quantum-interfering hard process & semi-classical decays: ~ no tempting partons!
- **Bringing truth tagging closer to reco**
  first releases used *b*-ancestry of jet constituents to set HF labels: too inclusive! ⇒ *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



# 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 ⇒ 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
  - 



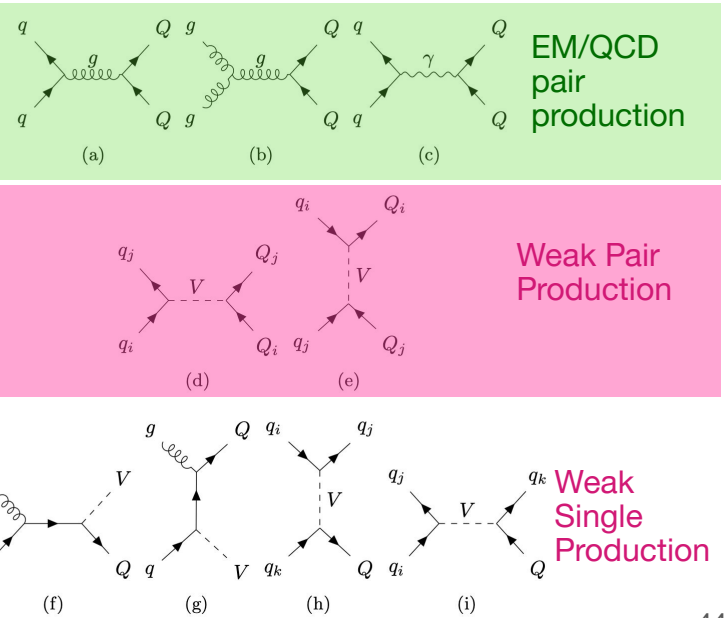
# Contur BSM example



❖ **Vector-like quarks** [SciPost Phys. 9, 069 (2020)]

- Popular generic class of SM extensions, with new quark partners: $B^{-1/3}$, $T^{2/3}$, $X^{5/3}$, $Y^{-4/3}$
- Couple to SM via usual quark EM & strong couplings, but
  *B,T*: interact with *W*, *Z* or *H* via modified weak coupling
  *X, Y*: interact only with *W* via modified weak coupling: $X \rightarrow Wt$, $Y \rightarrow Wb$ always
- LHC searches mostly for 3rd gen, strong pair-production only!
- 4 masses, 1 overall coupling κ,
  3 generational couplings ζ, 3 branching ratios ξ
  ⇒ rich collider phenomenology!

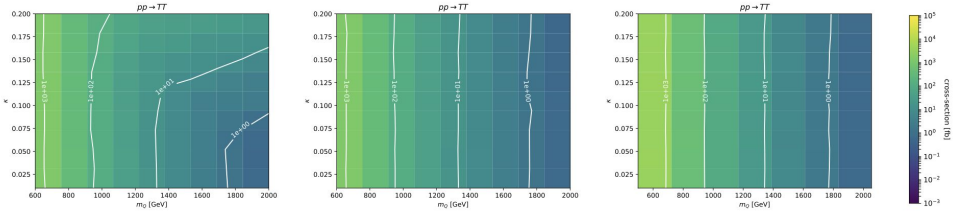
$$\mathcal{L} = \kappa_B \left[ \sqrt{\frac{\zeta_i \xi_W^B}{\Gamma_W^0}} \frac{g}{\sqrt{2}} [\bar{B}_{L/R} W_\mu^- \gamma^\mu u_{L/R}^i] + \sqrt{\frac{\zeta_i \xi_Z^B}{\Gamma_Z^0}} \frac{g}{2c_W} [\bar{B}_{L/R} Z_\mu \gamma^\mu d_{L/R}^i] - \sqrt{\frac{\zeta_i \xi_H^B}{\Gamma_H^0}} \frac{M_B}{v} [\bar{B}_{R/L} H d_{L/R}^i] \right]$$
$$+ \kappa_T \left[ \sqrt{\frac{\zeta_i \xi_W^T}{\Gamma_W^0}} \frac{g}{\sqrt{2}} [\bar{T}_{L/R} W_\mu^+ \gamma^\mu d_{L/R}^i] + \sqrt{\frac{\zeta_i \xi_Z^T}{\Gamma_Z^0}} \frac{g}{2c_W} [\bar{T}_{L/R} Z_\mu \gamma^\mu u_{L/R}^i] - \sqrt{\frac{\zeta_i \xi_H^T}{\Gamma_H^0}} \frac{M_T}{v} [\bar{T}_{R/L} H u_{L/R}^i] \right]$$
$$+ \kappa_X \left[ \sqrt{\frac{\zeta_i}{\Gamma_W^0}} \frac{g}{\sqrt{2}} [\bar{X}_{L/R} W_\mu^+ \gamma^\mu u_{L/R}^i] \right] + \kappa_Y \left[ \sqrt{\frac{\zeta_i}{\Gamma_W^0}} \frac{g}{\sqrt{2}} [\bar{Y}_{L/R} W_\mu^- \gamma^\mu d_{L/R}^i] \right] + \text{h.c.},$$



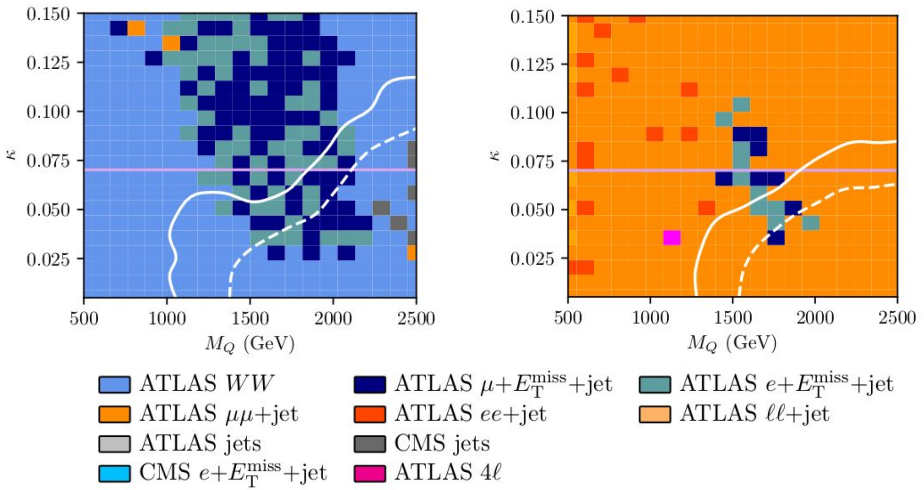
# VLQ pheno with Contur: 1st gen

- Even pair-production has $\kappa$-dependence via weak production initiated by valence quarks
- Weak *Qq* single-VLQ production can dominate over pair-prod!
- Different *W:Z:H* BFs for *T, B* activate different analysis pools “automatically” due to Rivet coverage
- *WW* diboson mostly dominates, thanks to *W* and *H* decay channels
- “Injection” of ℓ+MET+jet events here from an unfolded VBF *control region*!

Pair-prod cross-sections



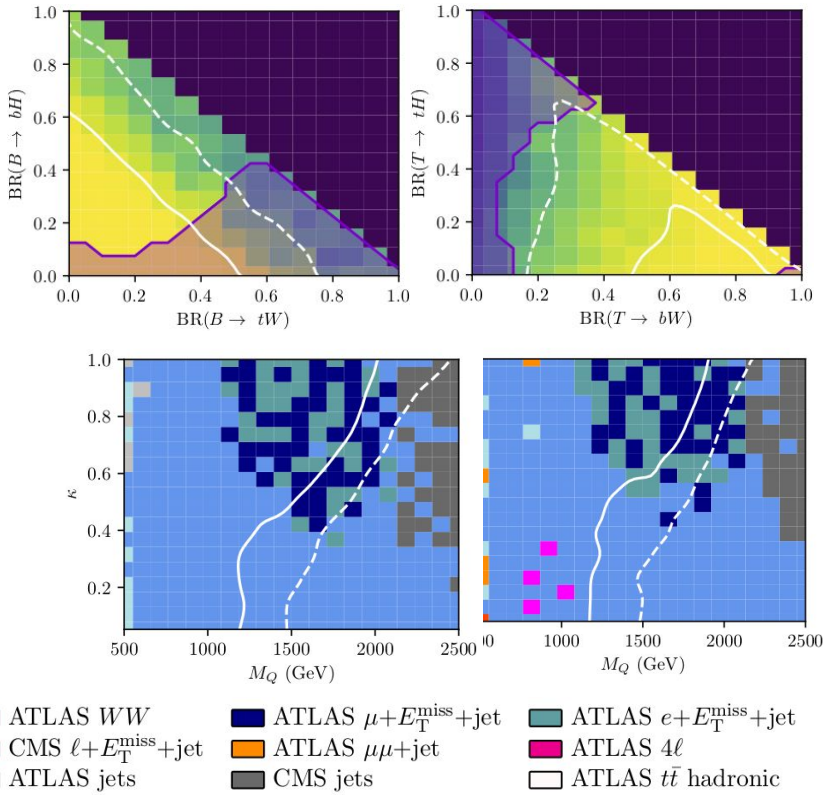

(a) $T$ coupling to $u, d$ only (b) $T$ coupling to $c, s$ only (c) $T$ coupling to $t, b$ only




Exclusions complementary to non-collider limits

# VLQ pheno with Contur: 3rd gen

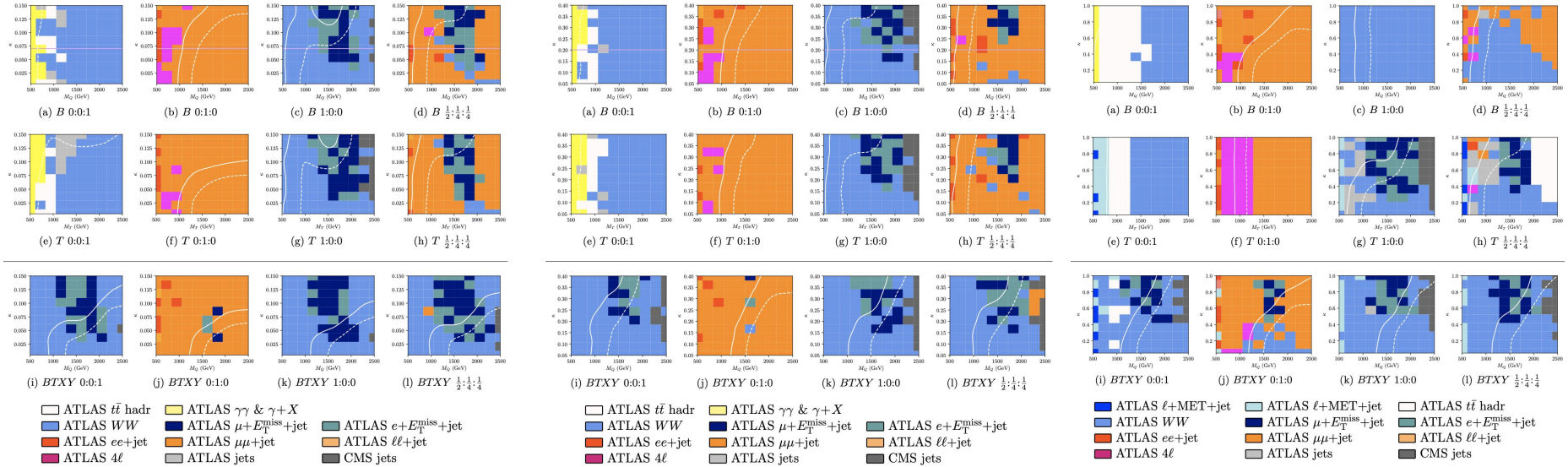
- In pure T, B pair-production mode, diboson and ℓ+MET+jet "SM" analyses ~cover or complement direct searches wonderfully
- In general, for *W:Z:H* = 0:1:0, *Tq* and *Xq* production killed by tiny top-quark PDF: pairs at low-$m_Q$, *Yq* at high-mass. Decays always have a *W* (directly or via $T \to t\,Z$) ⇒ℓ+MET pool always dominates
- Rivet+Contur "SM" routines give powerful sensitivity to VLQs, even far from the benchmark search modes




Generalising to 4 VLQs, still strong exclusions

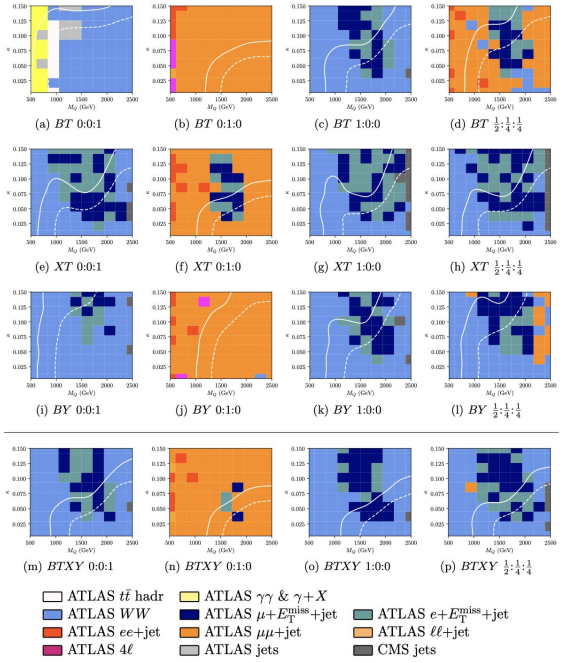
# More realistic models… [singlets]

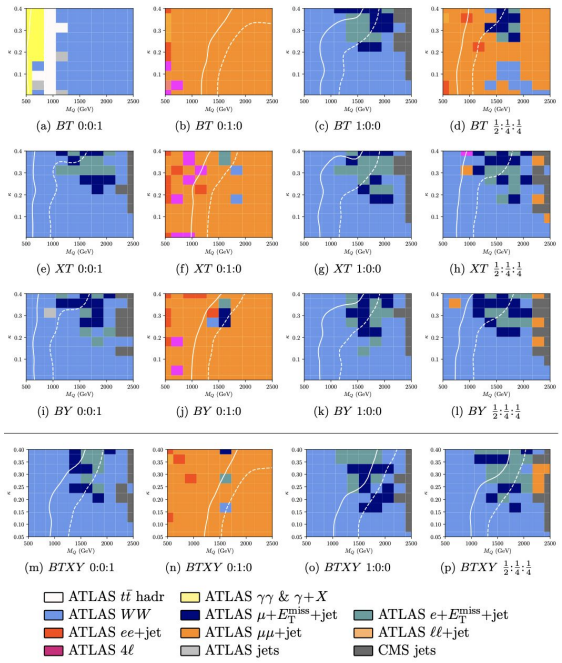
- Review requested a scan of realistic multiplets: 7 multiplets, each with 3 generational couplings, each with 4 W/H/Z-couplings, 300 points per scan, x 30,000 events!
- **~No problem! 1 month later…**

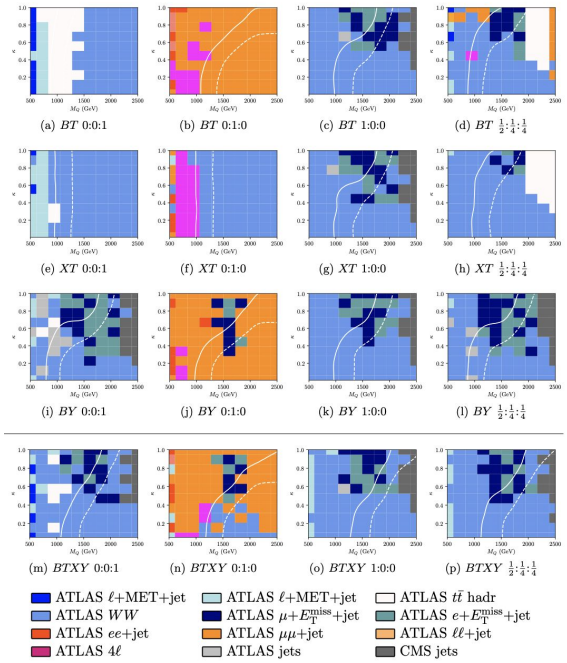


# More realistic models… [doublets]

- 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!

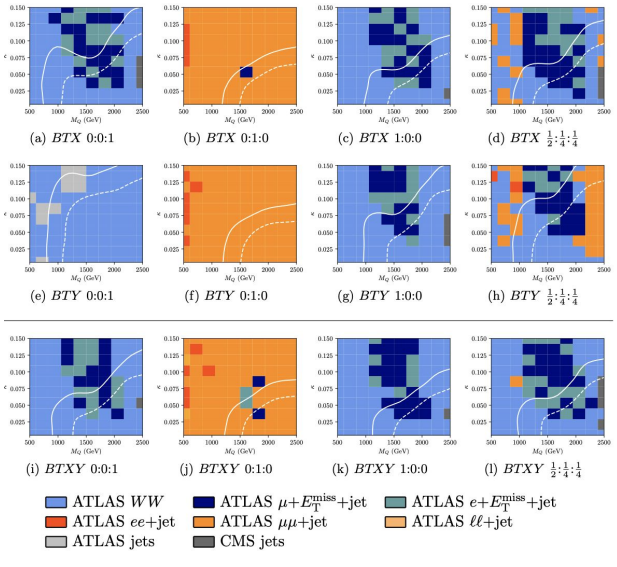


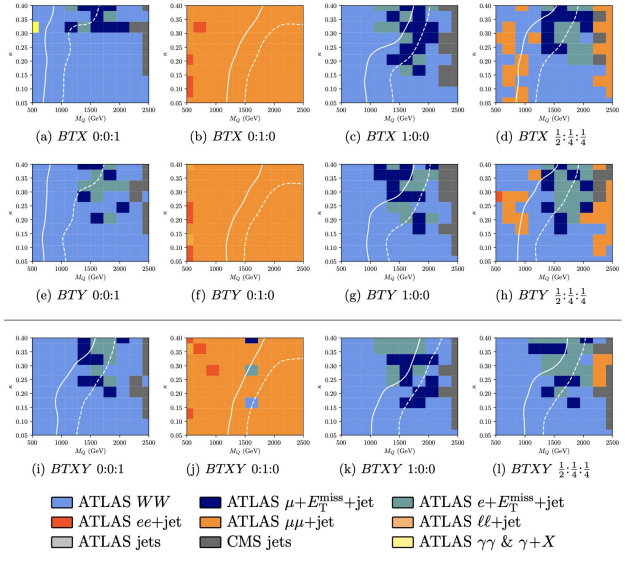



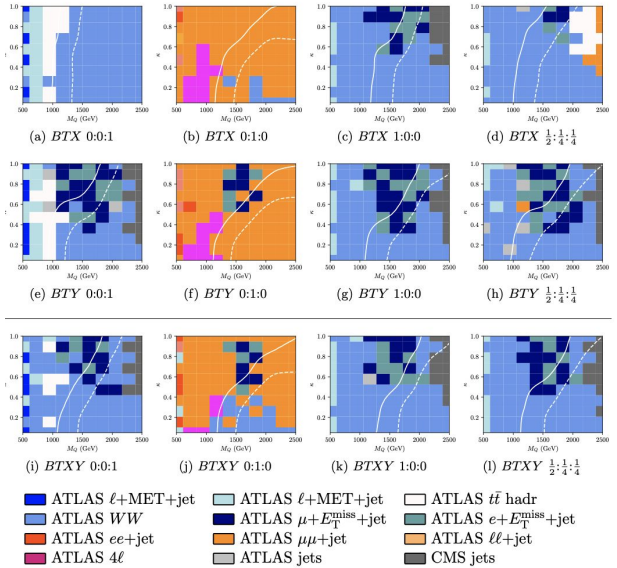



# More realistic models… [triplets]

- 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!




**Speed is *good!***

# Harder, faster, stronger… moar BSM

- Now extending beyond 1D and 2D grids:
  - ➢ Rivet (and Herwig) as a function
  - ➢ Embed into adaptive scans
  - ➢ Higher param dimensionalities
  - ➢ Including beyond colliders, e.g. Gambit
- Rivet as a tool to probe new-observable sensitivity, e.g. in EFT models (TopFitter)
- Bootstrapping for victory: estimating statistical and systematic correlations (with SModelS, MadAnalysis5)