Measurement and Monte Carlo

or "how to make a *useful* measurement"

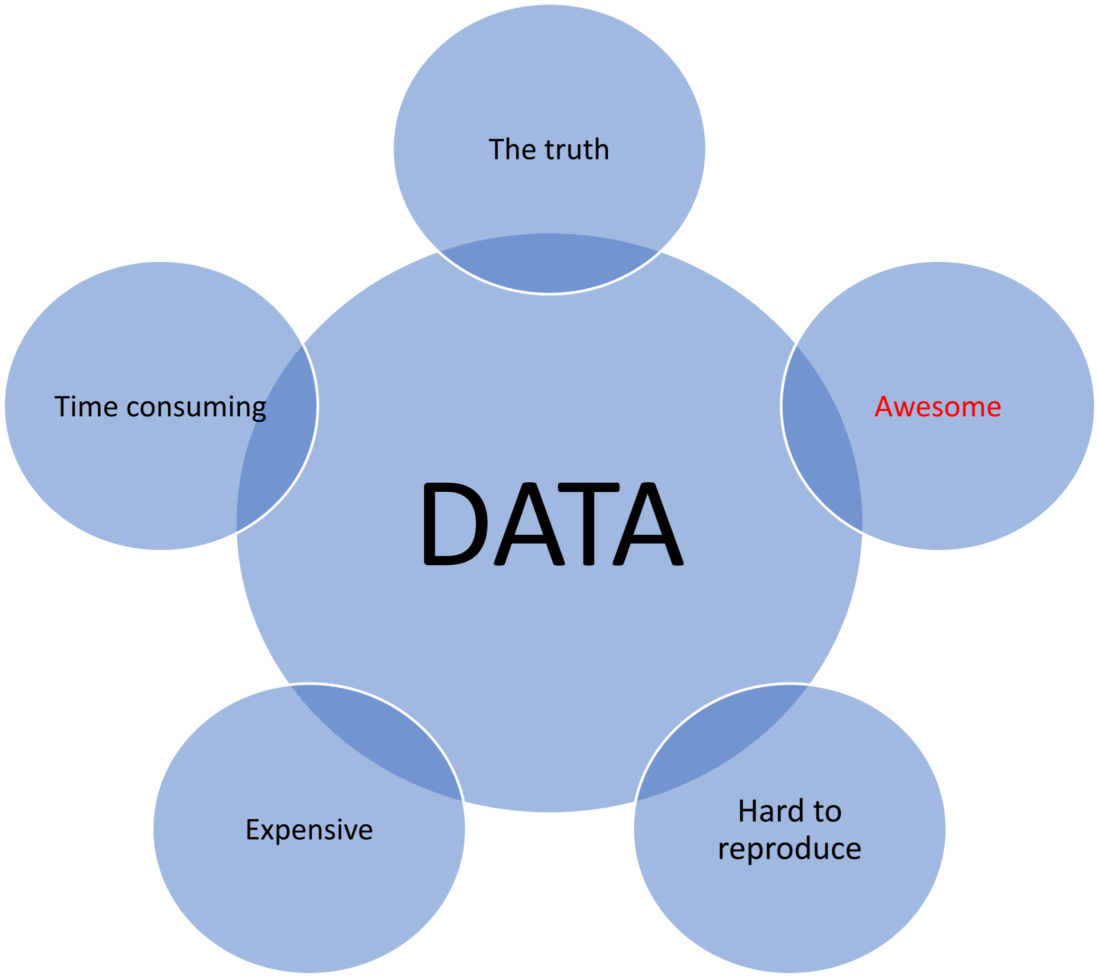
Part I

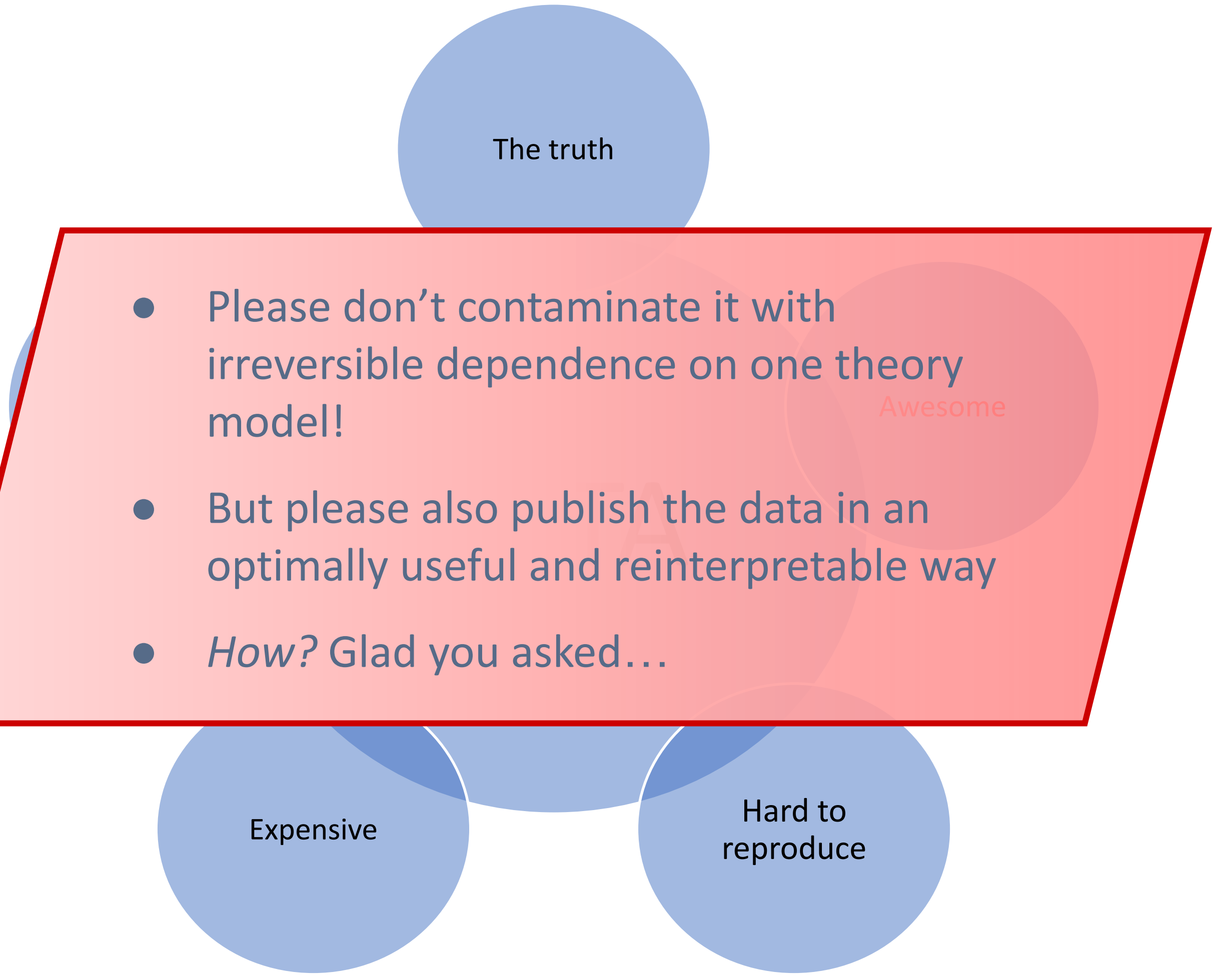
Andy Buckley

University of Glasgow

MCnet-Beijing Summer School, June 2021

*Thanks to Jon Butterworth and Emily Nurse for the use of their slides


The truth
Time consuming
Awesome
DATA
Expensive
Hard to reproduce



- Please don’t contaminate it with irreversible dependence on one theory model!
- But please also publish the data in an optimally useful and reinterpretable way
- *How?* Glad you asked…

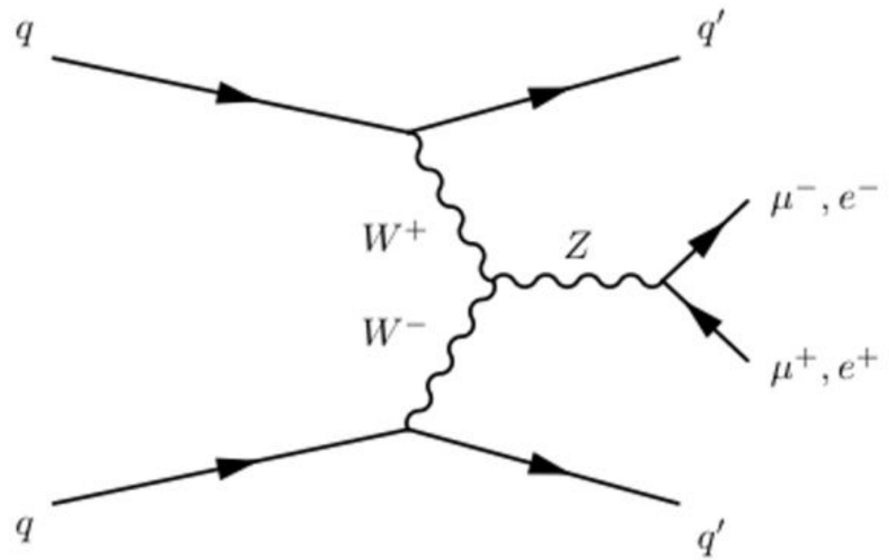
Overview

- What do we measure?
- Introduction to Monte Carlo generators in data-analysis
- Detector effects on various particles
- Making measurements as useful and model independent as possible:
  - Correcting for detector effects
  - (Not) extrapolating
  - The concept of a fiducial phase-space
  - What we mean by “final-state particles” (it is not always simple)
  - Background subtraction (or not)
- BSM measurements

This presentation is LHC-focused, and will have some bias towards ATLAS!
But all principles are applicable elsewhere

What do we actually measure?

- Electronic signals in detectors due to interactions with traversing particles produced in collisions
- Signals from multiple sub-detectors are combined and each collision “event” is reconstructed to give a list of identified particles/jets with kinematics



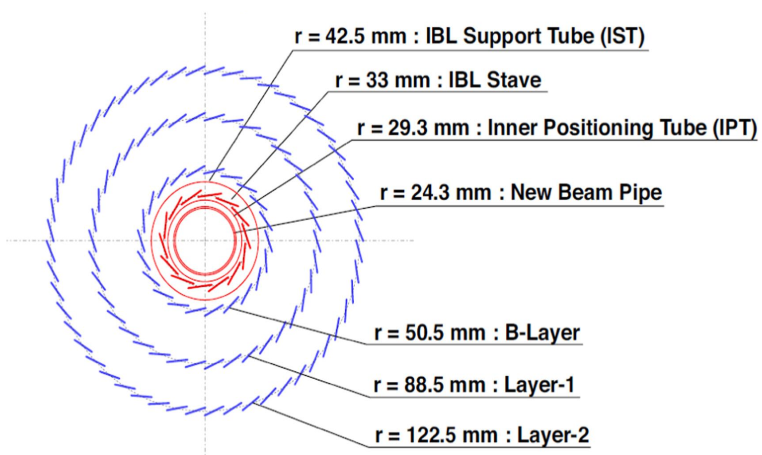
What do we actually measure?

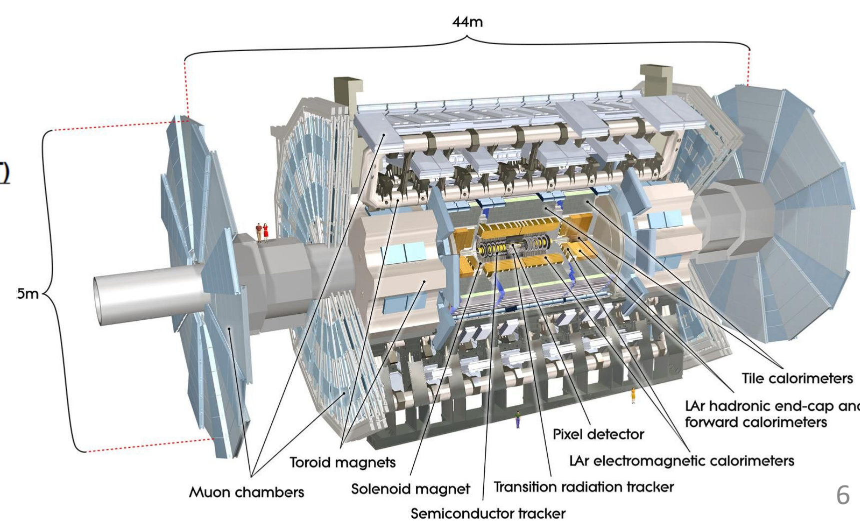
We only "see" *stable final-state* particles :

- electrons: stable
- muons: stable (τ_0 = 2.2 μs, mean decay after ~ 1 km at 20 GeV!!)
- taus: unstable (τ_0 = 0.3 ps, mean decay after ~ 1 mm at 20 GeV!!)
- neutrinos: stable (but invisible)
- photons: stable
- hadrons: unstable
  - ⟶ more leptons (charged and neutral), photons, more hadrons…
- quarks and gluons: unstable — are they even real??
  - ⟶ hadrons ⟶ jets
- W, Z, Higgs, top: unstable, varying degrees of objective reality!
  - ⟶ everything!

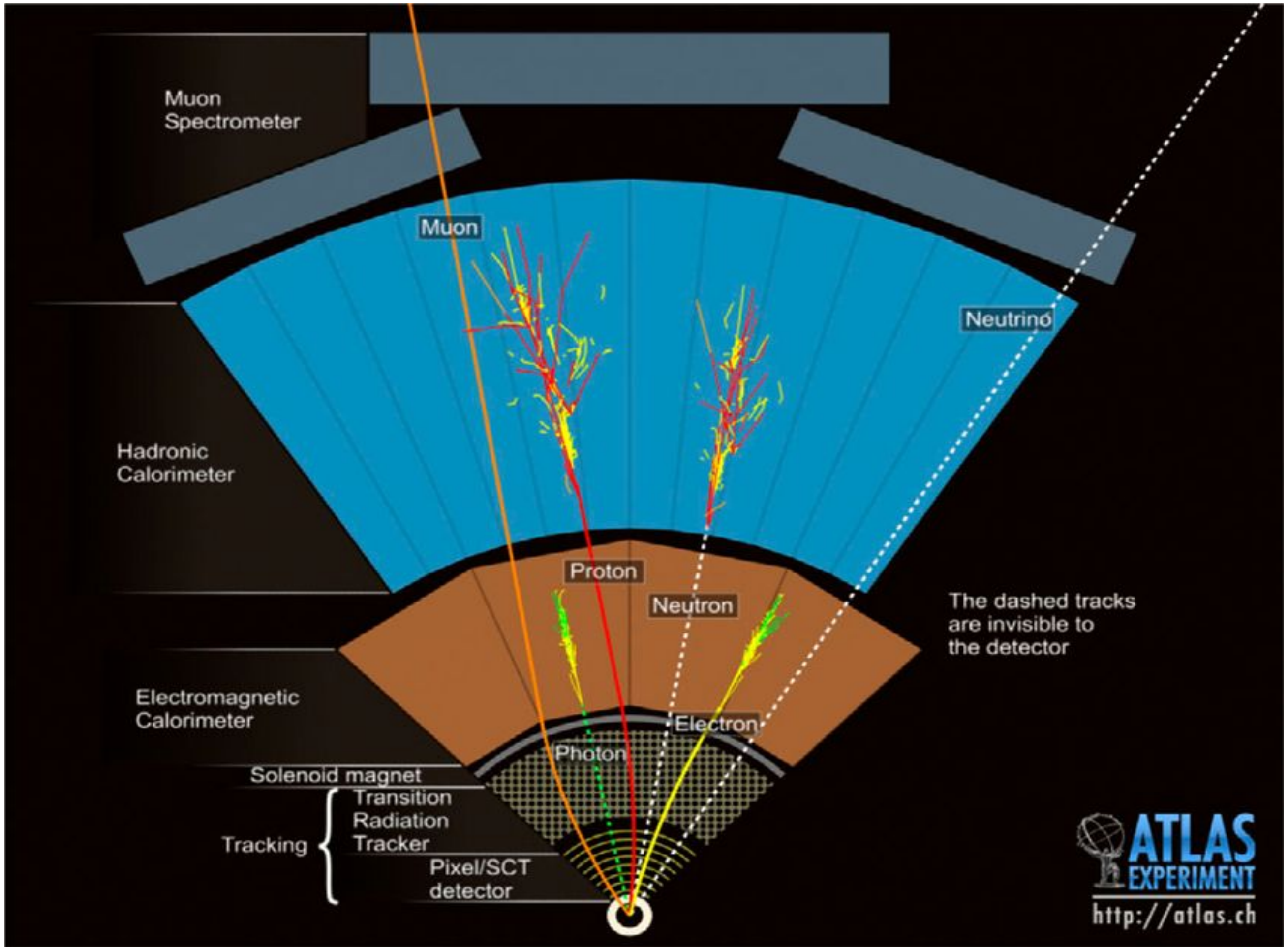
$$\boxed{d = \tau_0 \gamma v = c\tau_0 p \,/\, m}$$

$$m_\mu = 0.1\ GeV$$
$$m_\tau = 1.8\ GeV$$



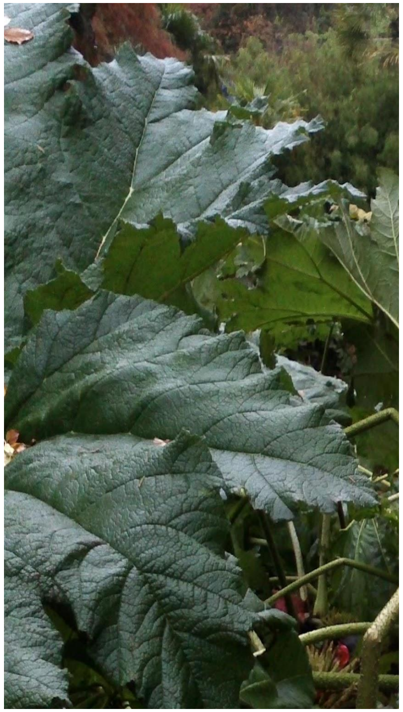


What do we actually measure?



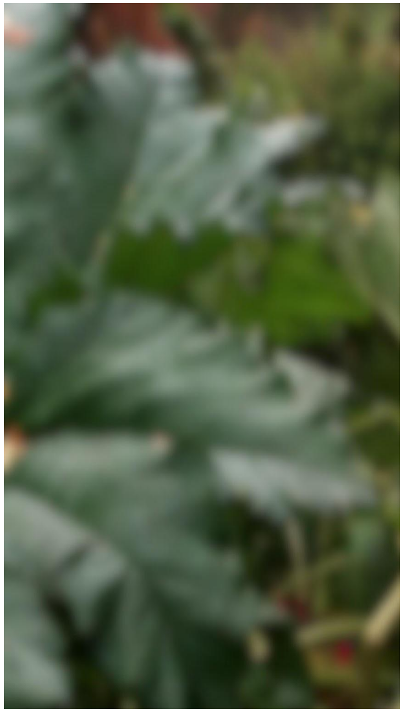
What do we actually measure?

- The kinematics of the identified particles are also reconstructed, and information about the event can be inferred



What do we actually measure?

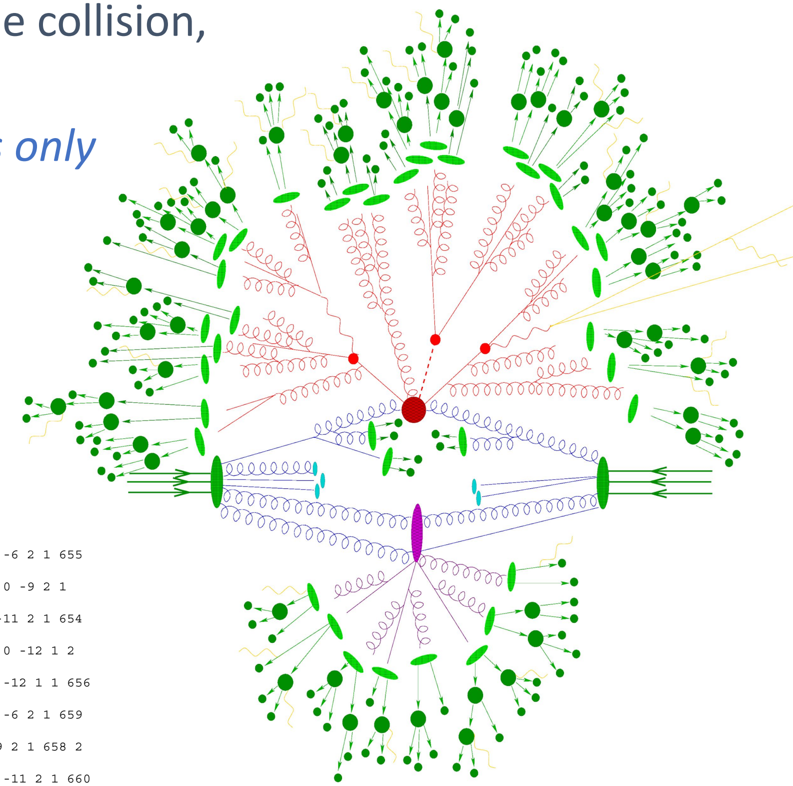
- The kinematics of the identified particles are also reconstructed, and information about the event can be inferred
- But these measurements are *not exact,* they have an *experimental resolution.* (And probably systematic biases, too.)



MC generators in measurement

- Event generators simulate collision events based on an underlying theory combined with phenomenological models with parameters tuned to experimental data (usually for low-energy QCD effects)
- The output is a list of particles produced in the collision, together with kinematics (four-vectors)
- *This part is experiment-independent, depends only on incoming particle types and CoM energy*

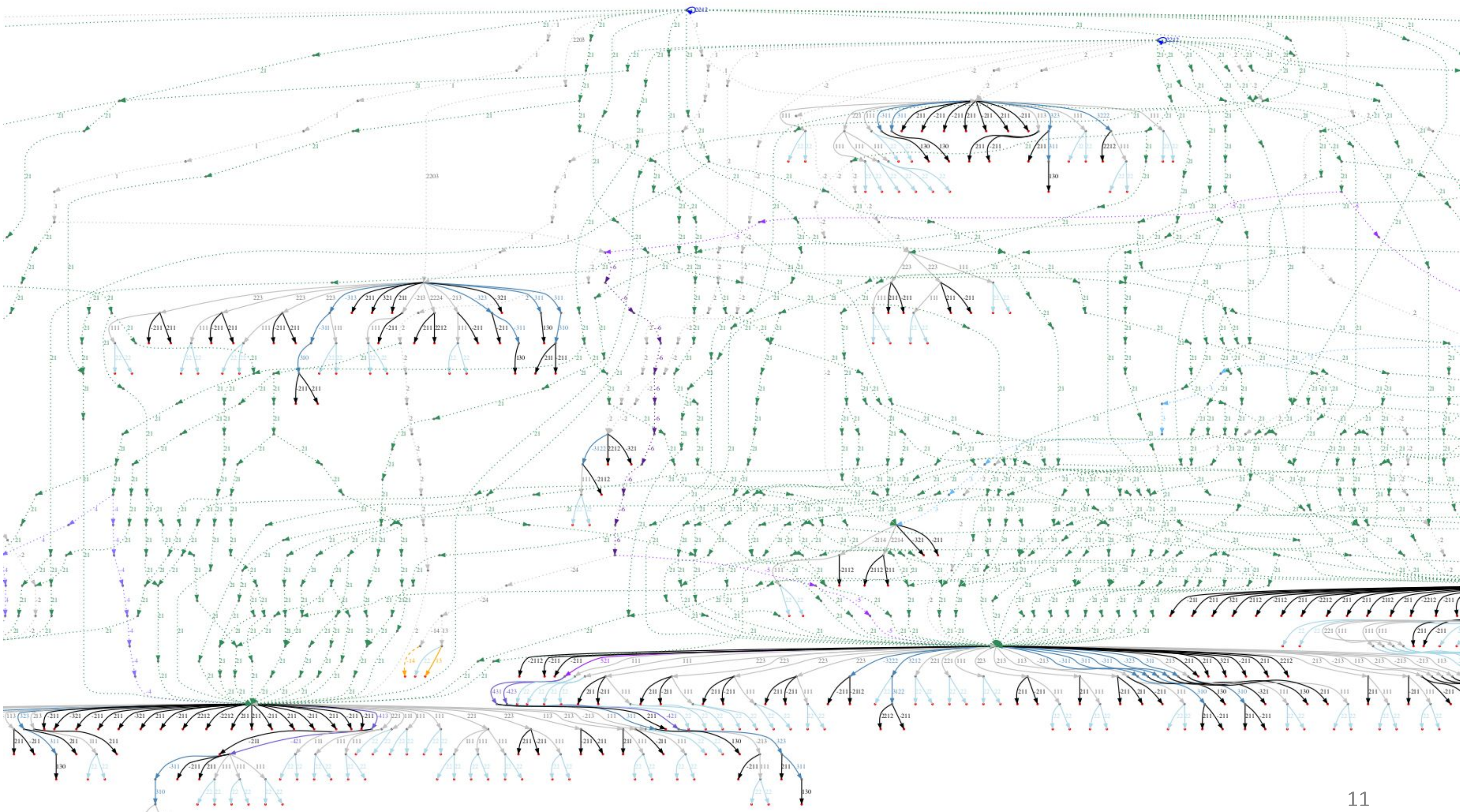

```
HepMC::Version 2.06.09
HepMC::IO_GenEvent-START_EVENT_LISTING
E 0 -1 0 1.305047132963e-01 7.763841138914e-03 0 -5 234 10001 10003 0 9 3.301267434432e-06 8.978821408834e+02
7.930514580328e+02 7.930514580328e+02 7.930514580328e+02 7.105872898865e+02 4.000000000000e+00 7.930514580328e+02
6.298240114645e+03
N 9 "MEWeight" "MUR0.5_MUF1_PDF261000" "MUR1_MUF0.5_PDF261000" "MUR1_MUF1_PDF261000" "MUR1_MUF2_PDF261000"
"MUR2_MUF1_PDF261000" "NTrials" "Weight" "WeightNormalisation"
U GEV MM
C 1.982628645082e+02 1.982628645082e+02
F 3 21 1.355269110210e-01 1.127542580157e-03 8.823075221978e+01 1.355269110210e-01 2.792889203654e+01 0 0
V -1 0 0 0 0 0 1 1 1 1.000000000000e+00
P 10001 2212 0 0 6.499999932280e+03 6.500000000000e+03 9.382720033633e-01 4 0 0 -1 0
P 10002 2212 0 0 6.499999932280e+03 6.500000000000e+03 9.382720033633e-01 11 0 0 -4 0
V -2 0 0 0 0 0 1 1 1 1.000000000000e+00
P 10003 2212 0 0 -6.499999932280e+03 6.500000000000e+03 9.382720033633e-01 4 0 0 -2 0
P 10004 2212 0 0 -6.499999932280e+03 6.500000000000e+03 9.382720033633e-01 11 0 0 -3 0
V -3 0 0 0 0 0 0 5 1 1.000000000000e+00
P 10005 21 1.714330700467e+00 2.213281091146e-01 -9.575581739813e+02 9.575597341546e+02 -2.157918643758e-05 11 0 0 -6 2 1 655
2 656
P 10006 21 -1.757323314213e+00 -4.154628631199e+00 -4.924799664895e+01 4.945416360863e+01 -1.383649647574e-05 11 0 0 -9 2 1
657 2 654
P 10007 21 1.582987254987e+00 2.799715977806e+00 -2.760412681726e+02 2.760600043333e+02 3.814697265625e-06 11 0 0 -11 2 1 654
2 655
P 10008 2101 -1.321999312907e+00 1.020529656020e+00 -3.814444371002e+03 3.814444780601e+03 5.793299988339e-01 11 0 0 -12 1 2
657
P 10009 2 -2.179953283341e-01 1.130548882582e-01 -1.402590739555e+03 1.402590761053e+03 -1.525878906250e-05 11 0 0 -12 1 1 656
V -4 0 0 0 0 0 0 5 1 1.000000000000e+00
P 10010 21 -1.776658431622e+00 2.479865383302e-01 9.401401408359e+02 9.401418522880e+02 -1.078959321879e-05 11 0 0 -6 2 1 659
2 661
P 10011 21 1.999658988953e+00 8.983465456712e-01 1.336251894549e+03 1.336253692735e+03 3.051757812500e-05 11 0 0 -9 2 1 658 2
659
P 10012 21 -1.730206524559e+00 6.026174210027e-02 4.297680545482e+02 4.297715415849e+02 -8.374976501503e-05 11 0 0 -11 2 1 660
2 658
P 10013 2203 1.736227309883e+00 -1.470428846972e+00 3.155057560022e+03 3.155058474679e+03 7.713299971049e-01 11 0 0 -12 1 2
660
P 10014 1 -2.290213426542e-01 2.638340208699e-01 6.386648994053e+02 6.386649949634e+02 7.629394531250e-06 11 0 0 -12 1 1 661
```

Picture from Sherpa authors

MC generators in measurement

MC event-record graphs are only partially physical! Which bits are safe?!

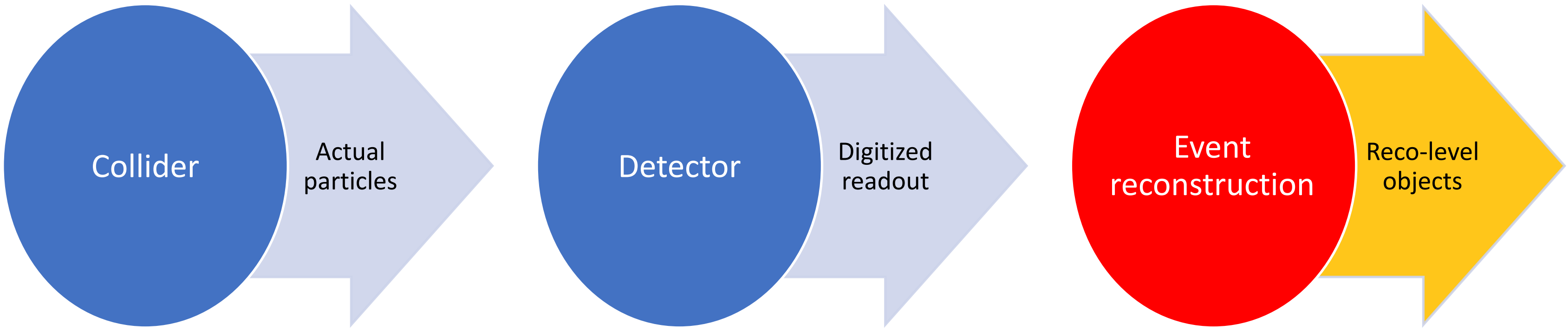


MC detector-simulation

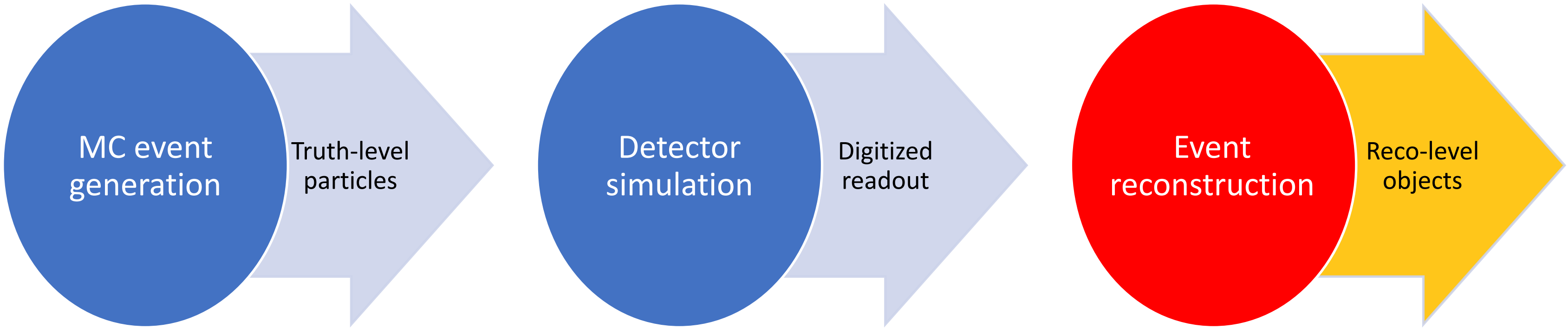
- Often we also have to simulate the effect of our detectors
- Special simulation codes based (usually) on Geant4 (and increasingly also custom codes to speed it up)
- Generated particles pass step-by-step through material (with which they interact) and magnetic fields (where they curve and radiate)
- Digitization step simulates detector response in terms of electronic signals (same format as data)
- *The same reconstruction code as used in data can then be applied to the simulated events*

- *This part is usually experiment-specific:* detector simulation is *CPU-intensive* and codes are *not publicly available*

“Real data” processing chain



MC-simulation processing chain



MC generators in data-analysis

Generated events are used to:

1. Compare measured data to expectations from a given theory (SM or otherwise). Usually we ask "does the data agree with this theory?"

2. Subtract expected background processes from the data *(we'll later discuss why this isn't always the best idea)*

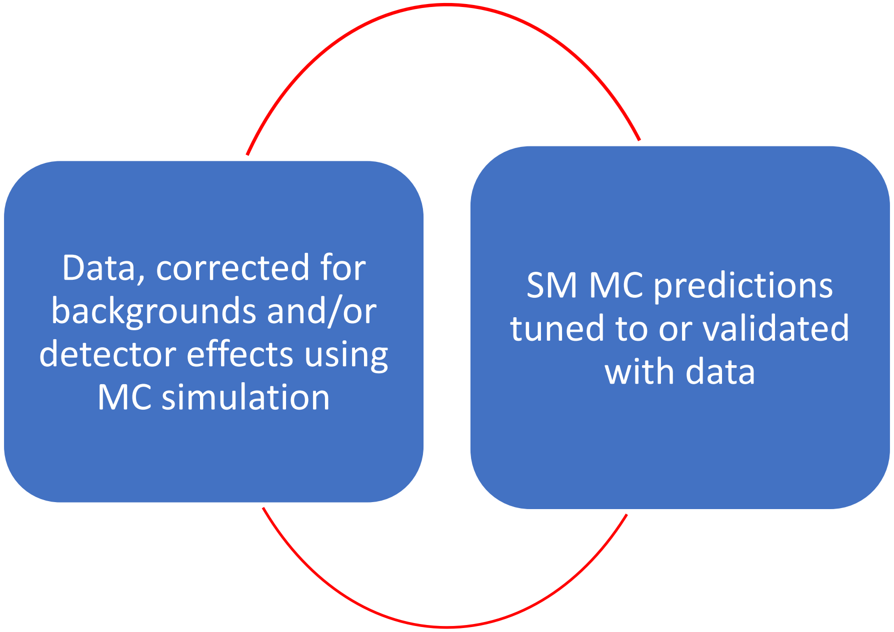
3. Correct for detector effects by comparing *truth-level* MC prediction with *reco-level* MC prediction *(more on this later)*

4. Plan the sensitivity of future experiments

For this it is often necessary to correct to correct for detector effects and present the data in terms of "truth-level" particles/objects

What do theorists *want*? ;-)

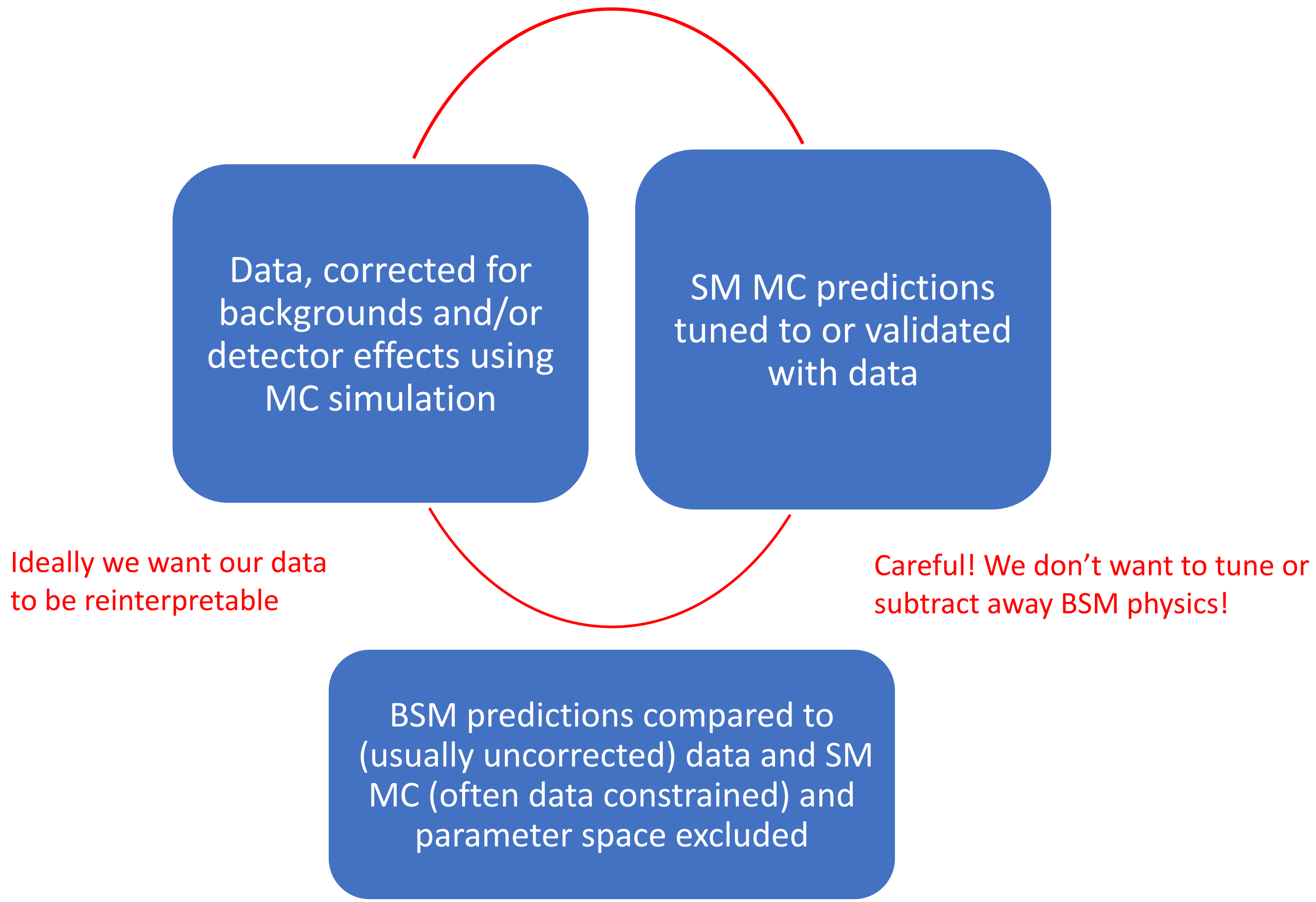
Usually they ask “How well does the data agree with my prediction?”
(the prediction often comes as a set of final-state
 “truth” particles from MC generation)



Careful! We don’t want the data to depend on the prediction we are constraining!

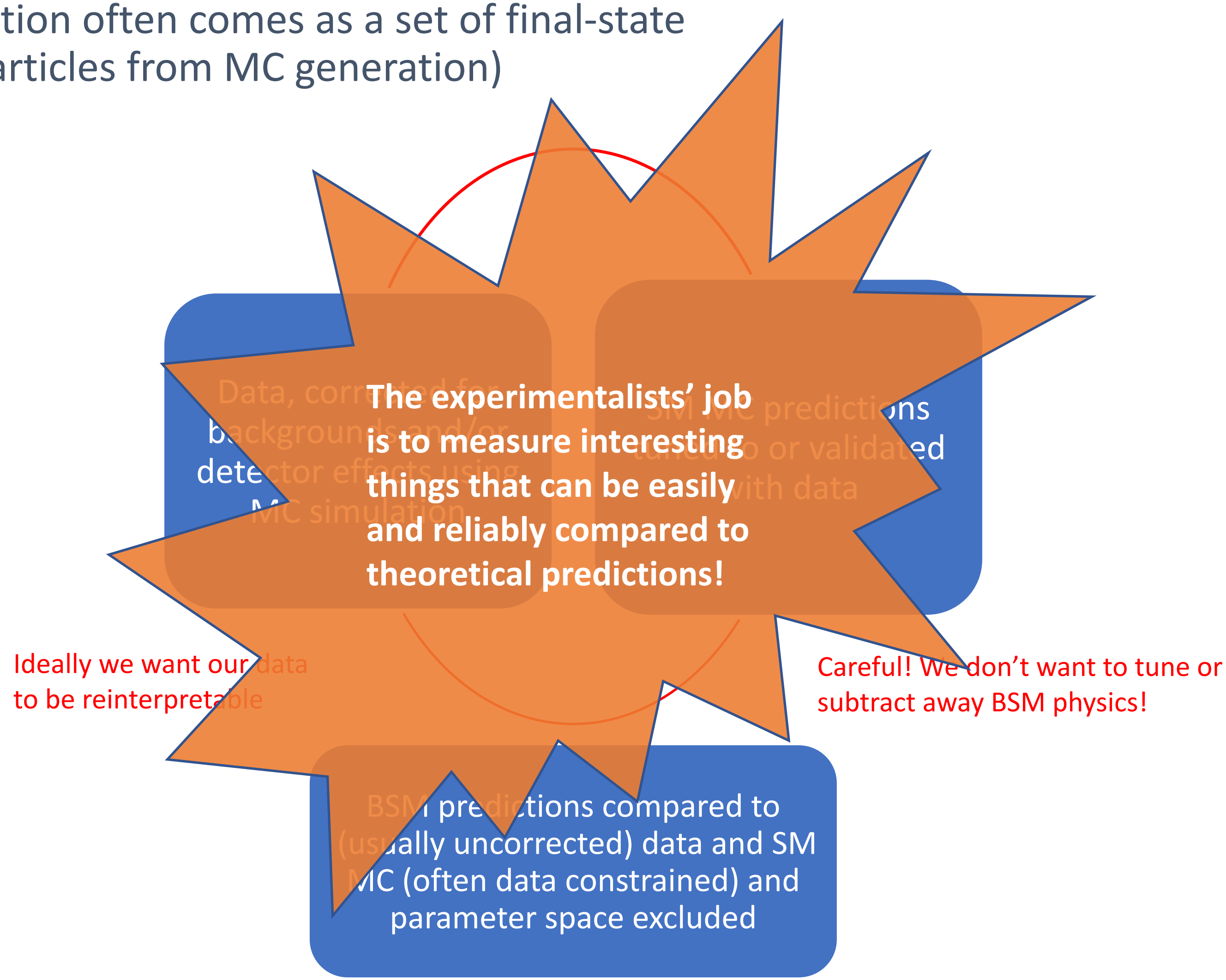
What do theorists *want*? ;-)

Usually they ask “How well does the data agree with my prediction?”
(the prediction often comes as a set of final-state “truth” particles from MC generation)



What do theorists *want*? ;-)

Usually they ask “How well does the data agree with my prediction?”
(the prediction often comes as a set of final-state
“truth” particles from MC generation)



Rivet and analysis-preservation

See the tutorials!
https://rivet.hepforge.org

- A system for validation of Monte Carlo event generators.
- Experimental results are included via HepData and an analysis routine is written that selects events and plots the relevant variables to compare to the data.
- Makes sure theorists are making the correct selection cuts when comparing to your data!! > 1000 analyses preserved so far
- Incredibly useful for MC generator development, validation, and tuning, as well as testing BSM-physics models

When you publish a result, please make sure you provide a Rivet routine, too!

Rivet analysis coverage

Rivet analyses exist for 324/5731 papers = 6%. 185 priority analyses required.

Total number of Inspire papers scanned = 7216, at 2019-05-21

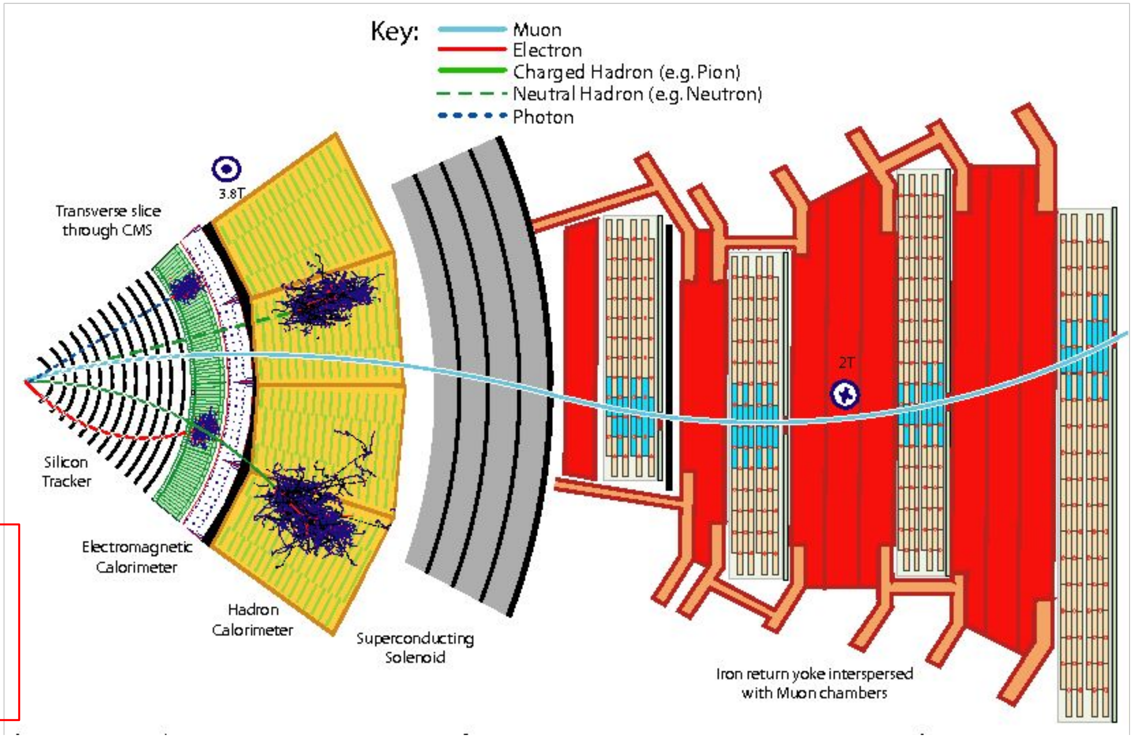
Breakdown by identified experiment (in development):

Key	ALICE	ATLAS	CMS	LHCb	B-factories	HERA	LEP	Other
Rivet wanted (total):	200	264	354	161	1498	446	1418	1066
Rivet REALLY wanted:	35	42	74	10	2	14	7	1
Rivet provided:	**20**/220 = **9%**	**149**/413 = **36%**	**77**/431 = **18%**	**11**/172 = **6%**	**14**/1512 = **1%**	**8**/454 = **2%**	**38**/1456 = **3%**	**7**/1073 = **1%**

https://rivet.hepforge.org/rivet-coverage

Detector effects and biases

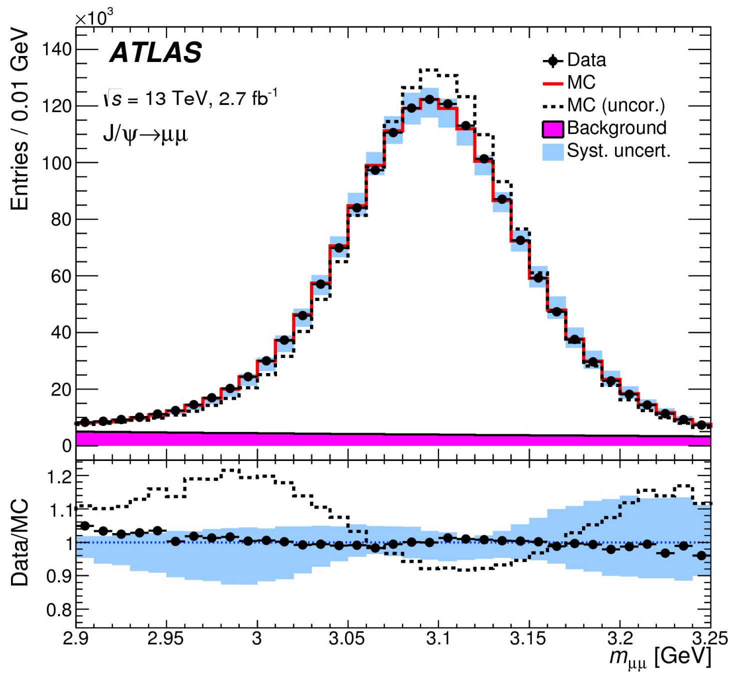
- **Efficiencies:** there is a non-zero probability that a particle passing through a detector will not be reconstructed
- **Fake backgrounds:** there is a non-zero probability that a particle will be reconstructed even though it wasn't really there
- **Smearing:** the measured energies, momenta, and angles of the particles and jets will be smeared due to the intrinsic resolutions of the detectors

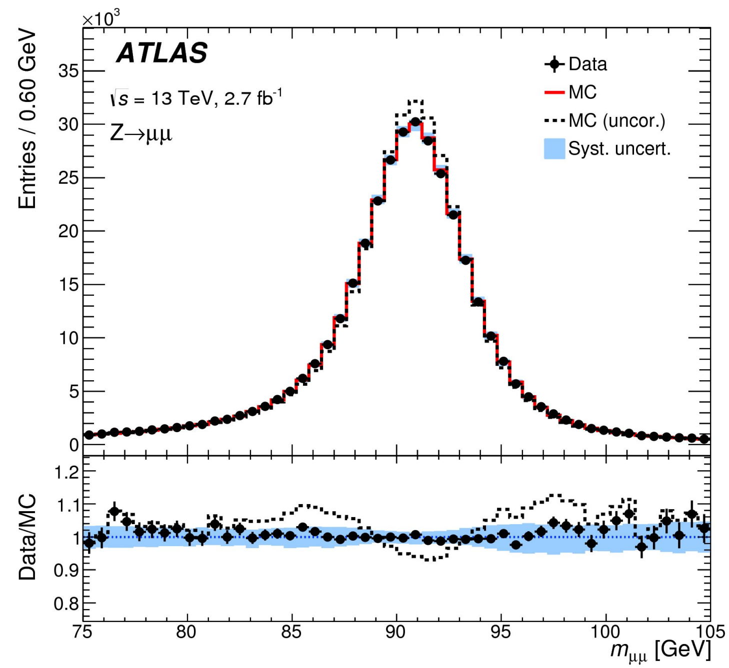


We need to know what our detector is doing so we can account for it — and in some cases reverse it

Detector effects: muons

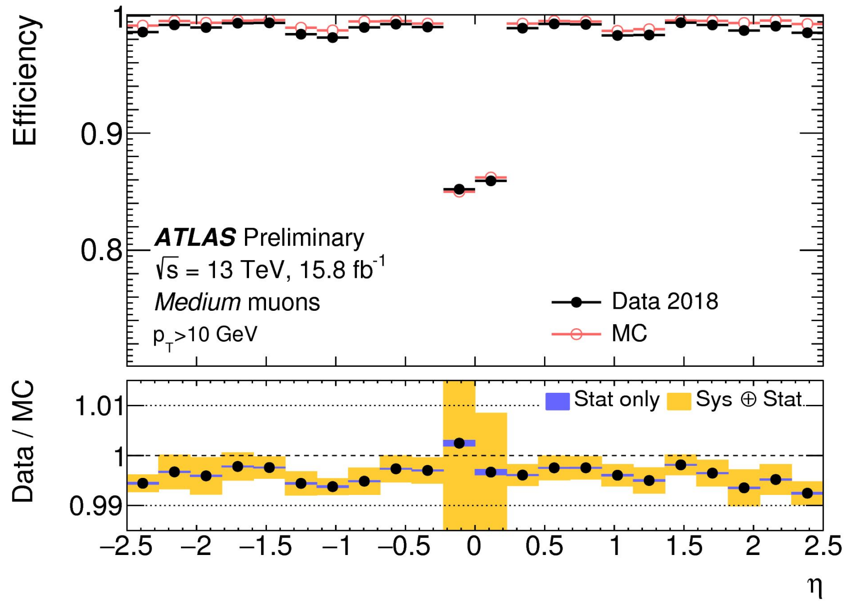
- Momentum measured in tracker and muon detectors via charged-particle tracks
- Usually with isolation requirements (keep away from electrons and jets, to reduce contamination of “signal” by hadron ⟶ μ + X decays
- Reconstruction calibrated via “standard candle” Z ⟶ μ μ and J/ψ ⟶ μ μ mass peaks

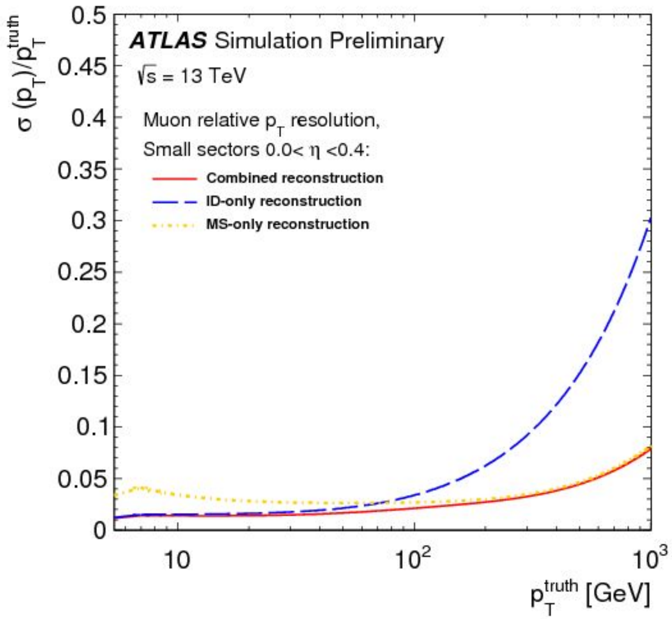




[Eur. Phys. J. C 76 (2016) 292]

Detector effects: muons

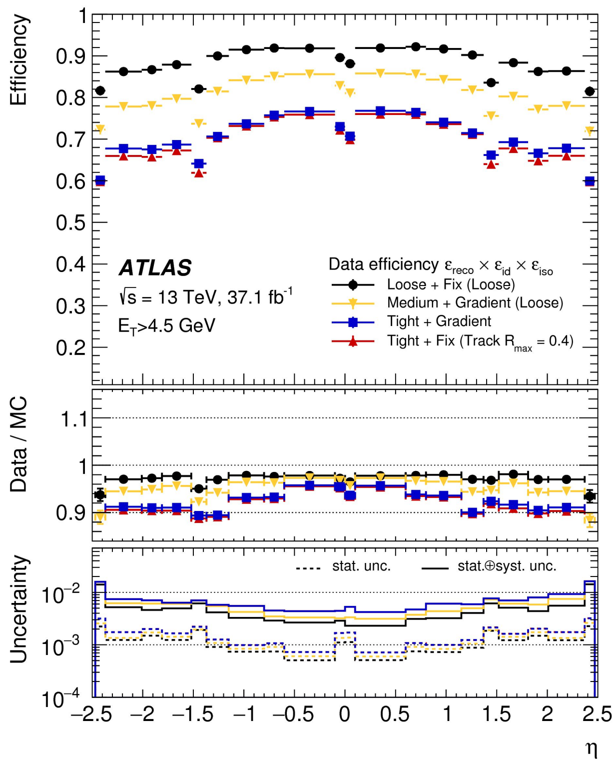


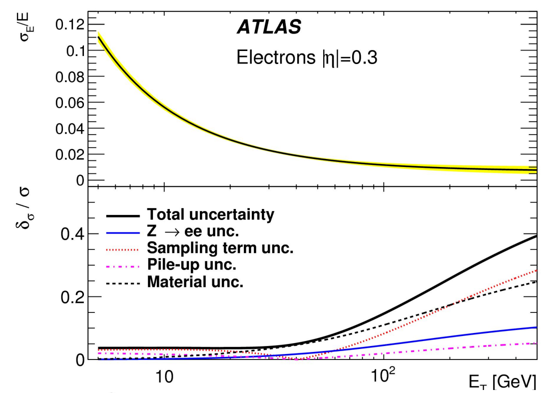


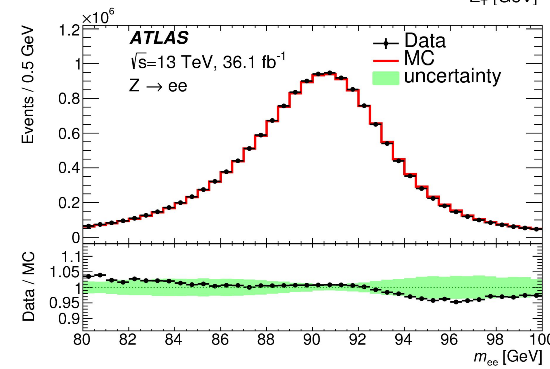
- High reconstruction efficiency
- Percent-level p_T resolution at low p_T (gets worse at high p_T — straight tracks!)

Detector effects: electrons and photons

- Calorimeter signal-cluster measures energy; electrons matched to inner-detector tracks, and discriminated via shower-shape variables
- Usually isolation requirements again, to cut out hadron decays, e.g. $\pi^0 \rightarrow \gamma\,\gamma$





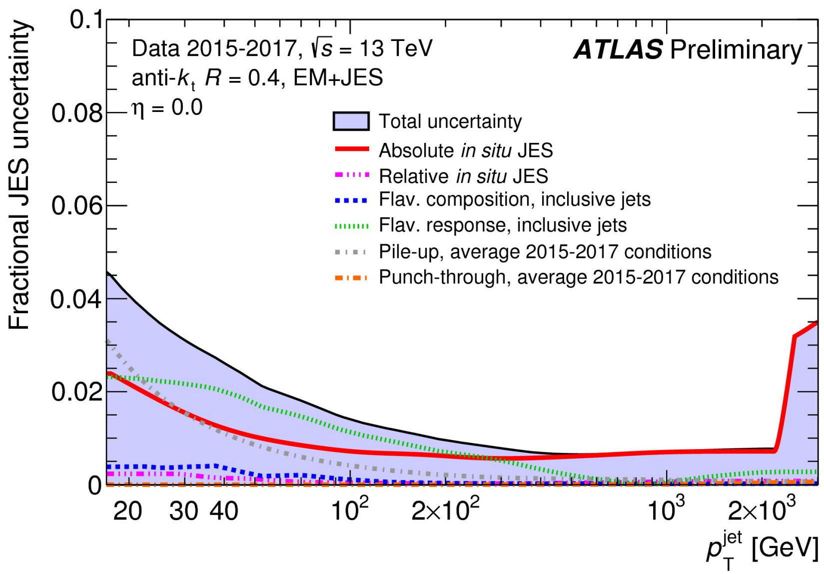


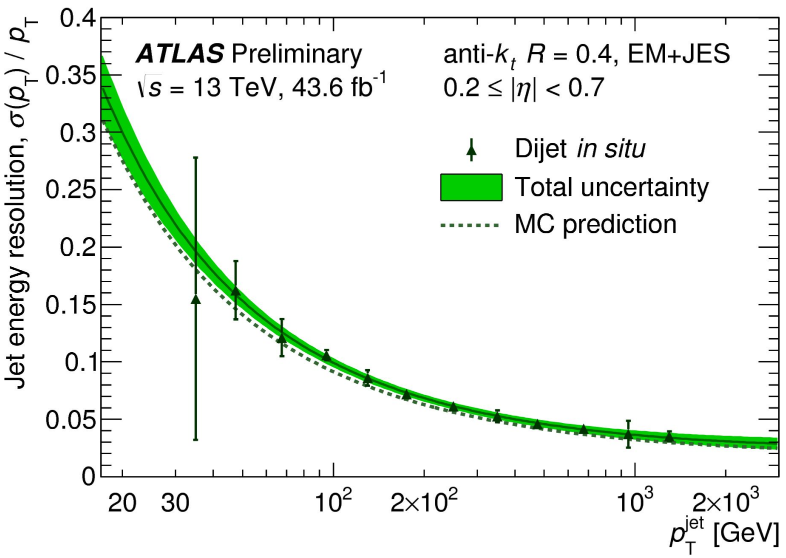
arXiv:1902.04655

- High reconstruction efficiency
- Energy resolution: percent-level at high-energy, gets worse at low energy (N_{clus})

Detector effects: jets

- Partons lead to collimated hadrons which we form into “jets”
- Built with jet algorithms (usually anti-k_T) from calorimeter clusters / tracks
- Calibrated by balance with other calibrated objects (electrons, muons, photons) and forward jets balanced with central jets

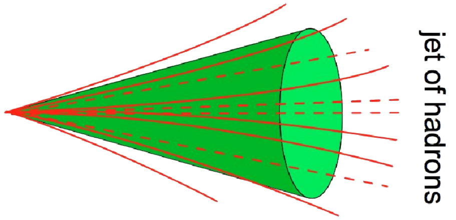


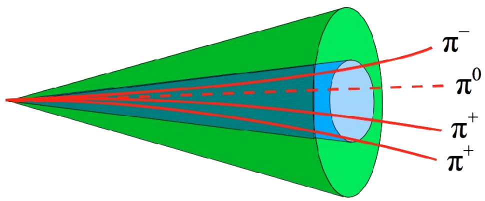


JETM-2018-006

Detector effects: (hadronic) taus

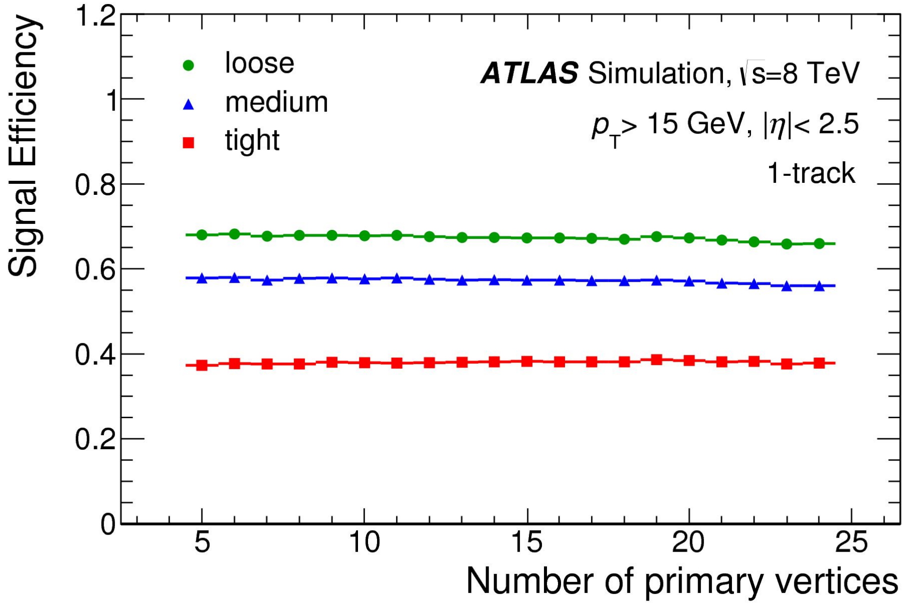
- Hadronic taus behave like hadronic jets
- But low and odd-number track multiplicity: identification by "prong counting"





Recall: τ_0 = 0.3 ps, decays after ~1 mm at 20 GeV

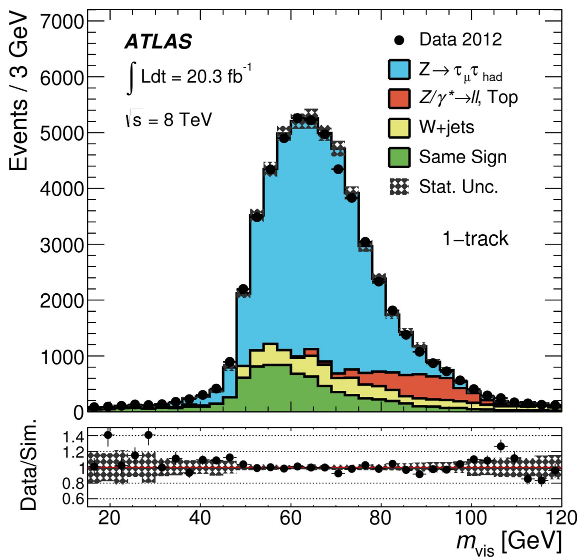
$$\tau^\pm \to \pi^\pm \pi^0 \nu$$
$$\tau^\pm \to \pi^\pm \nu$$
$$\tau^\pm \to \pi^\pm \pi^0 \pi^0 \nu$$
$$\tau^\pm \to \pi^\pm \pi^\pm \pi^\mp \nu$$
$$\tau^\pm \to \pi^\pm \pi^\pm \pi^\mp \pi^0 \nu$$

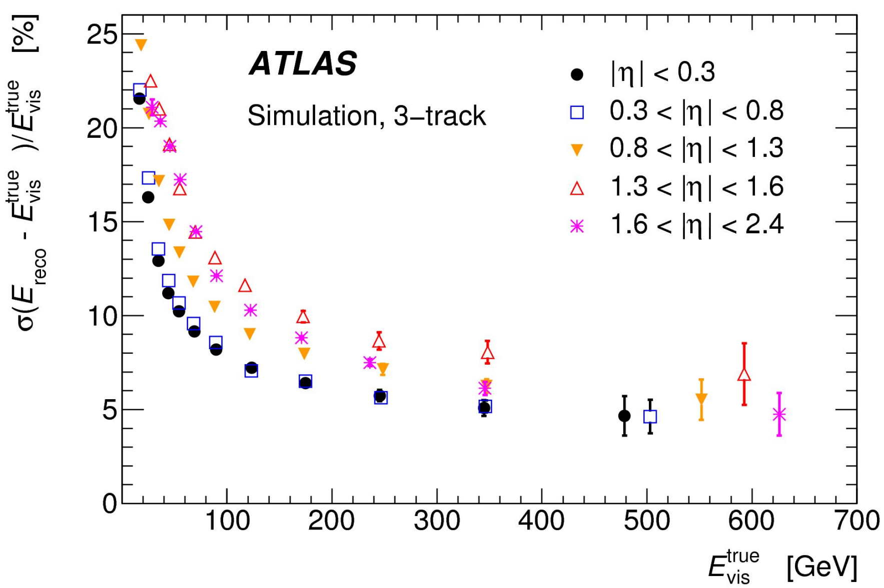


Eur. Phys. J. C75 (2015) 303

Detector effects: (hadronic) taus

- Calibrated to *visible* decay energy (i.e. not including neutrino)
- Resolution of 5—25% depending on E and η



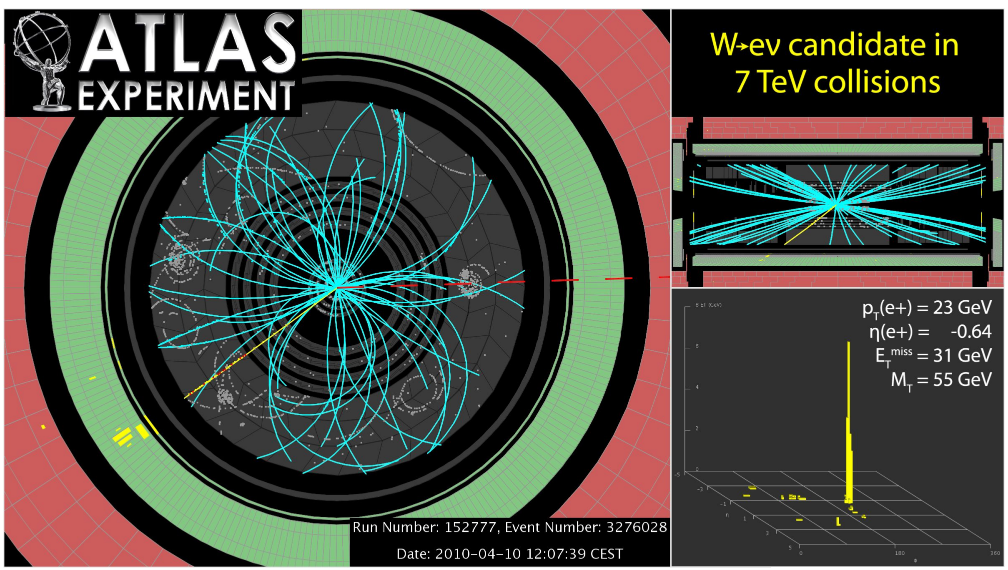


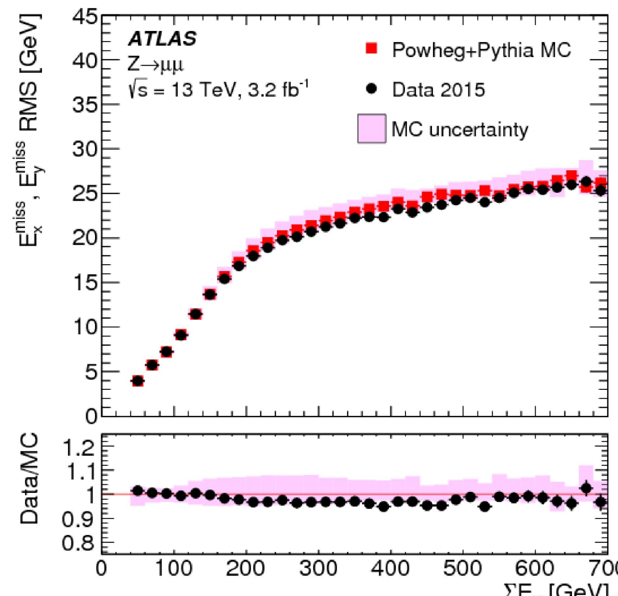
Eur. Phys. J. C75 (2015) 303

Detector effects: neutrinos etc. (aka MET or $E_{T,miss}$)

- Invisible particles — mainly neutrinos, but also e.g. BSM dark-matter candidates — aren't seen by the detector. Have to be inferred from absence of balance between the *visible* particles in the detector acceptance:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = \underbrace{\underbrace{-\sum_{\substack{\text{selected}\\\text{electrons}}} \mathbf{p}_{\mathrm{T}}^{e}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},e}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{photons}}} \mathbf{p}_{\mathrm{T}}^{\gamma}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\gamma}} \underbrace{-\sum_{\substack{\text{accepted}\\\tau\text{-leptons}}} \mathbf{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\tau_{\mathrm{had}}}} \underbrace{-\sum_{\substack{\text{selected}\\\text{muons}}} \mathbf{p}_{\mathrm{T}}^{\mu}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\mu}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{jets}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{jet}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\mathrm{jet}}}}_{\text{hard term}} \underbrace{\underbrace{-\sum_{\substack{\text{unused}\\\text{tracks}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{track}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\mathrm{soft}}}}_{\text{soft term}}$$





Recall: we often want to present the data corrected for detector effects so we can compare to final-state “truth-level” particles.

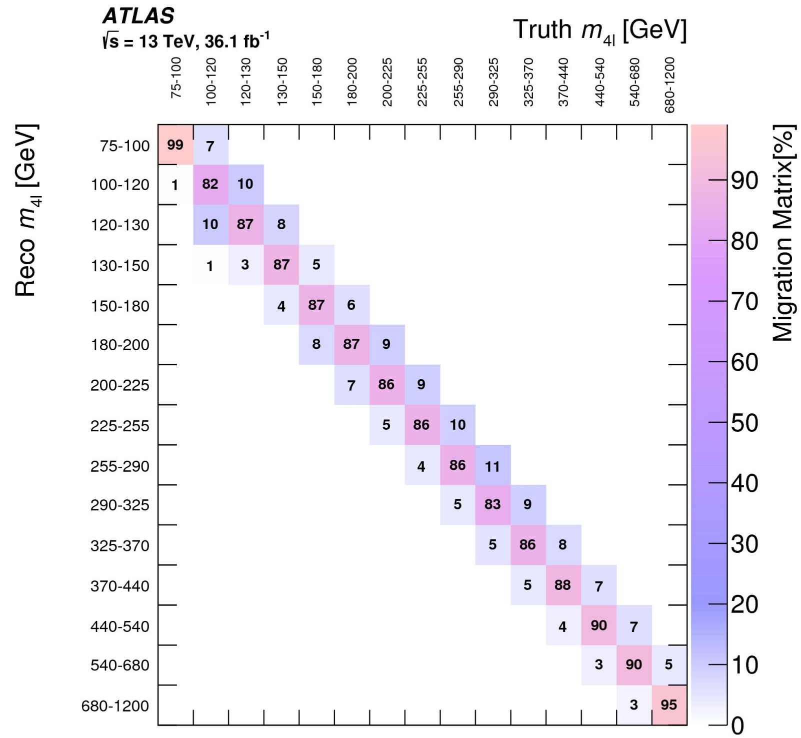
People outside the collaboration do not have access to CPU-intensive detailed detector-simulation codes with full detector geometry (and the full object-reconstruction framework to match)

Correcting for detector effects

- ✔ Correct for backgrounds from fake particles and sometimes those with similar final states
  (we will discuss later what to do with backgrounds leading to the *same* final state as the signal)

- ✔ Correct for the detector inefficiencies and scales and "unfold" resolution effects

- ✔ Assign systematic uncertainties to the corrected data to account for how well we understand the detector corrections



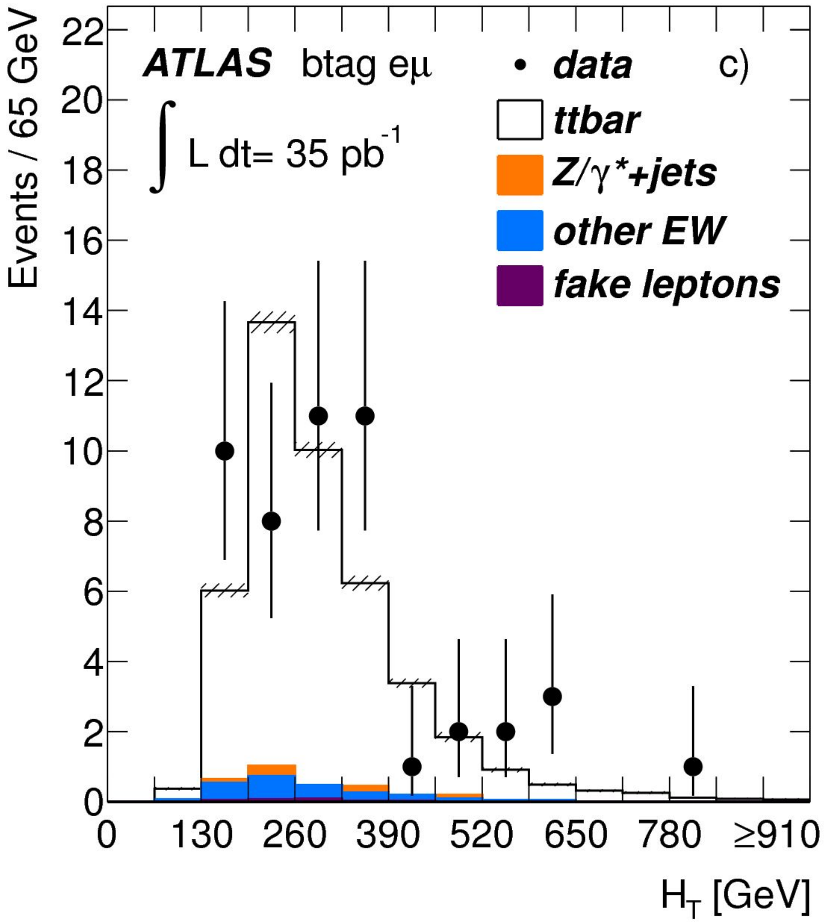
Only experimentalists can do this, and so they should! Only they know the details. Otherwise it is very hard to (re)-interpret an experimental result

Correcting for detector effects

- Corrections are derived using MC generators
- Need to account for instabilities and correlations in the corrections: "unfolding" statistical frameworks can either try to invert detector effects, or to fit the truth-level distribution values via many forward-foldings
- We must be careful as the corrections can depend on the underlying physics-modelling, e.g.
  - Bin migrations depend on underlying distribution
  - Efficiency corrections depend on kinematics of particles

- Validate / reweight underlying distributions by comparisons to data and assign appropriate systematic uncertainties
- Treat "MC A versus MC B" systematic uncertainties with caution

Uncorrected distributions



- Run 1 tt cross-section paper
- H_T distribution at *reco level*
- **This cannot be compared to any model prediction other than the one used in the paper**

Phys. Lett. B707 (2012) 459

Correcting for acceptance effects

- AKA extrapolating outside the measured region, into full phase-space
  - e.g. $p_T > 25$ GeV $\longrightarrow$ $p_T < 0$ GeV, and/or $|\eta| < 2.5 \longrightarrow |\eta| < \infty$
- Anyone can do this with a preferred SM prediction: no detector-simulation needed! Hurrah! *But hang on…*
- We didn't measure this region. We're inserting a 100% model-dependent prediction into the measurement. It is a bad idea to contaminate our very precious data with the very theory we are trying to constrain
  - At best, do an extrapolation, e.g. for comparison to non-MC theory, *in addition* to the "real" measurement
- An example:

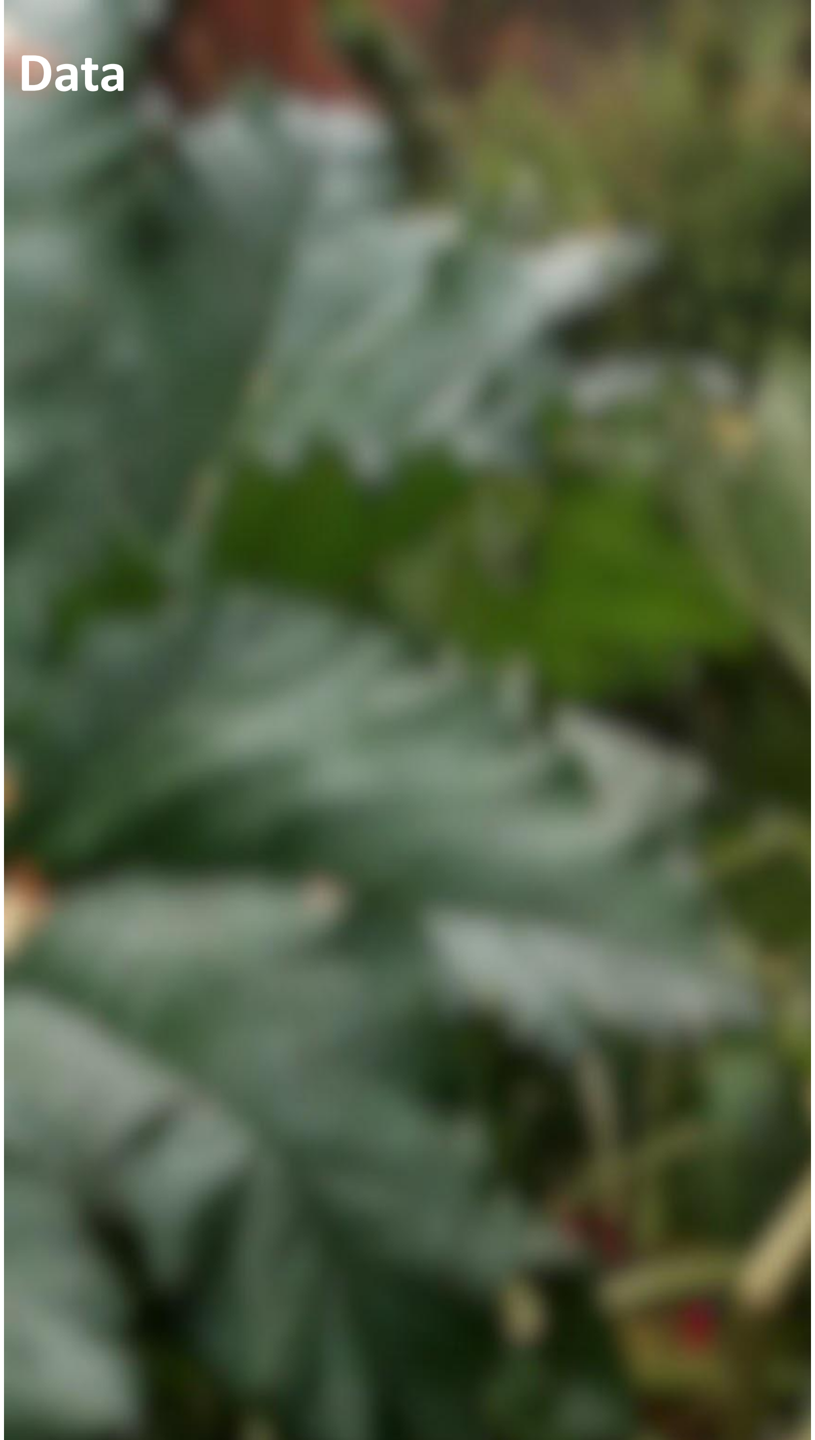
$$\boxed{\sigma_{t\bar{t}} = 177 \pm 25\, pb}$$

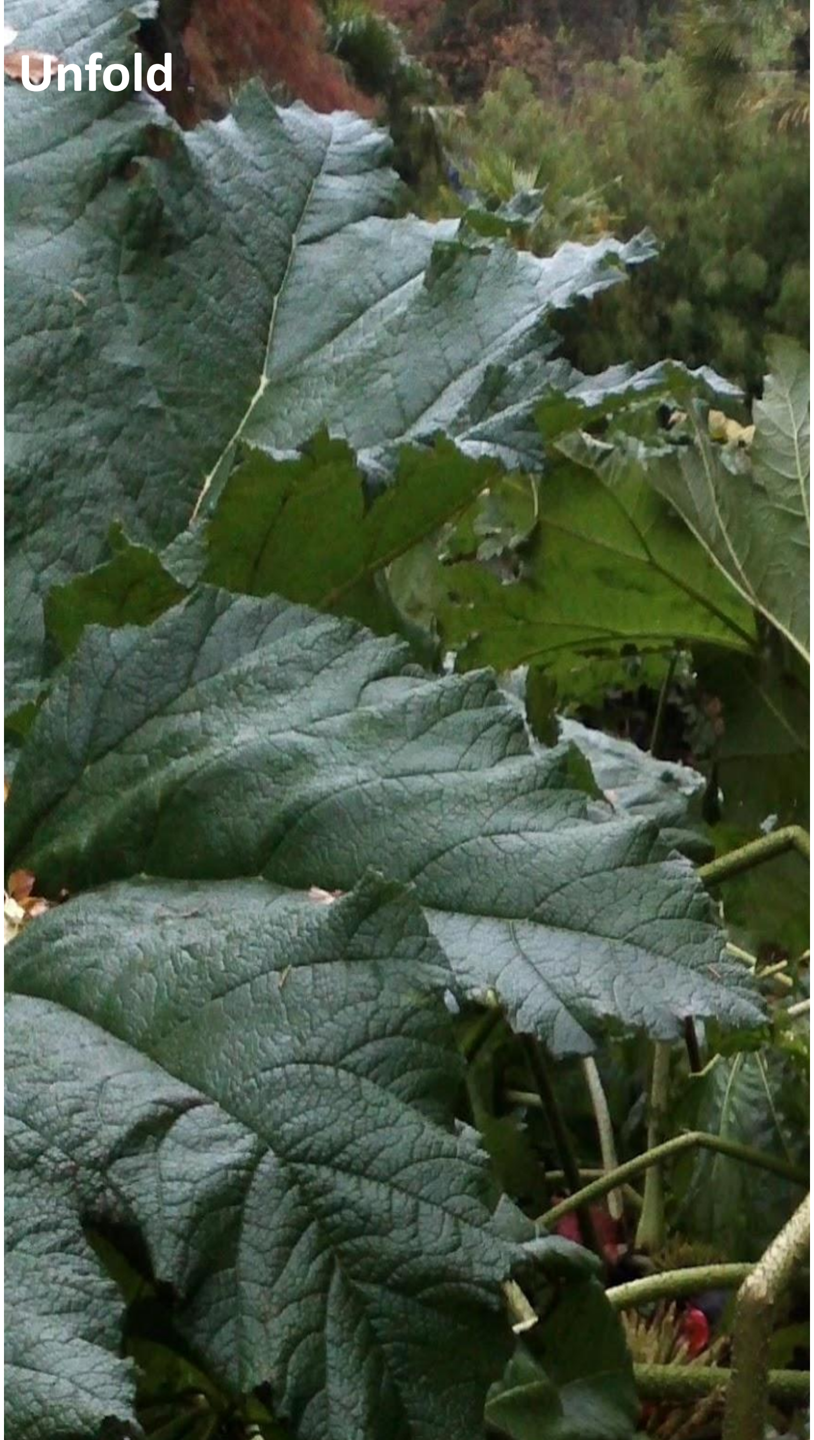
Phys. Lett. B707 (2012) 459

LHC Run 1 total tt cross-section, from a measurement made in the dilepton decay channel with $p_T < 25$ GeV, $|\eta| < 2.5$, and more cuts on $E_{T,miss}$, H_T, jets, etc.!

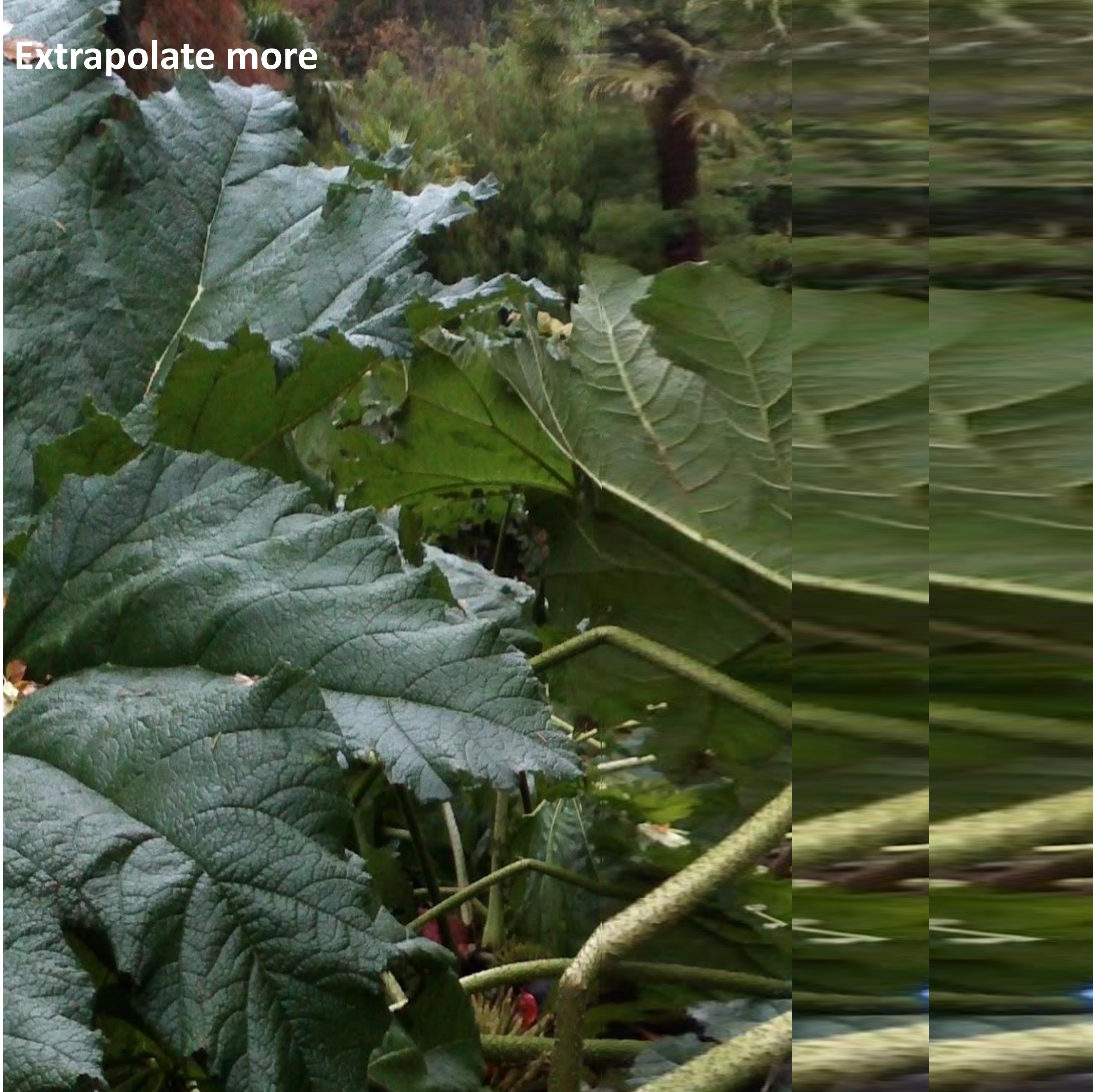
- Only 1.7% of tt events were used to measure the tt cross-section!! 98.3% of events were not seen. Some is from detector inefficiencies, but much was extrapolation into an unseen region… *don't do this!*

Data




Unfold

Extrapolate


Extrapolate more

But how reliably?

Avoiding dodgy extrapolations motivates the idea of a "fiducial measurement" …

More on that in Part 2!