Measurement and Monte Carlo

or "how to make a *useful* measurement"

Part II

Andy Buckley

University of Glasgow

MCnet-Beijing Summer School, June 2021

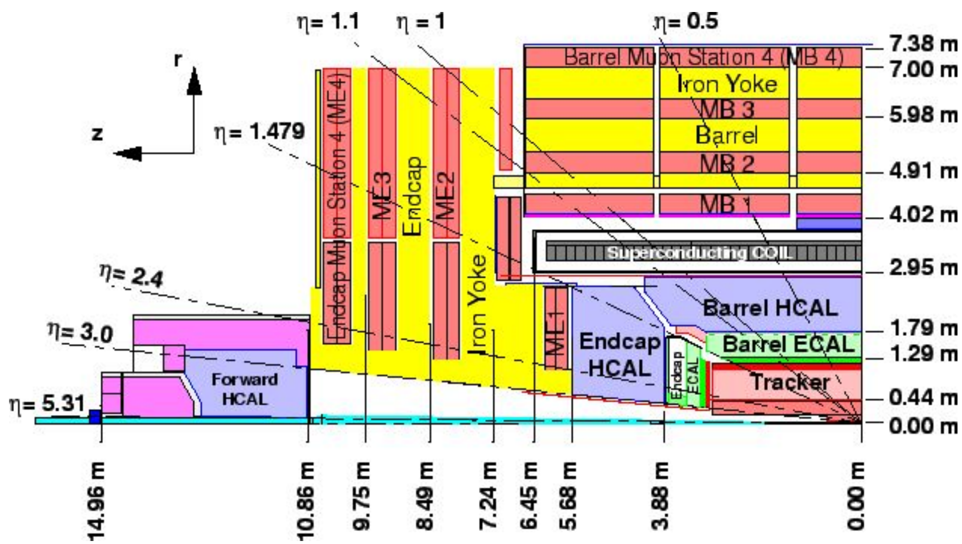
*Thanks to Jon Butterworth and Emily Nurse for the use of their slides

Fiducial phase-space

Regardless of detector efficiencies and resolutions, there are uninstrumented kinematic regions that we don't measure *at all*:

We do not have 4π detectors, and can't reconstruct particles down to zero pT!

CMS



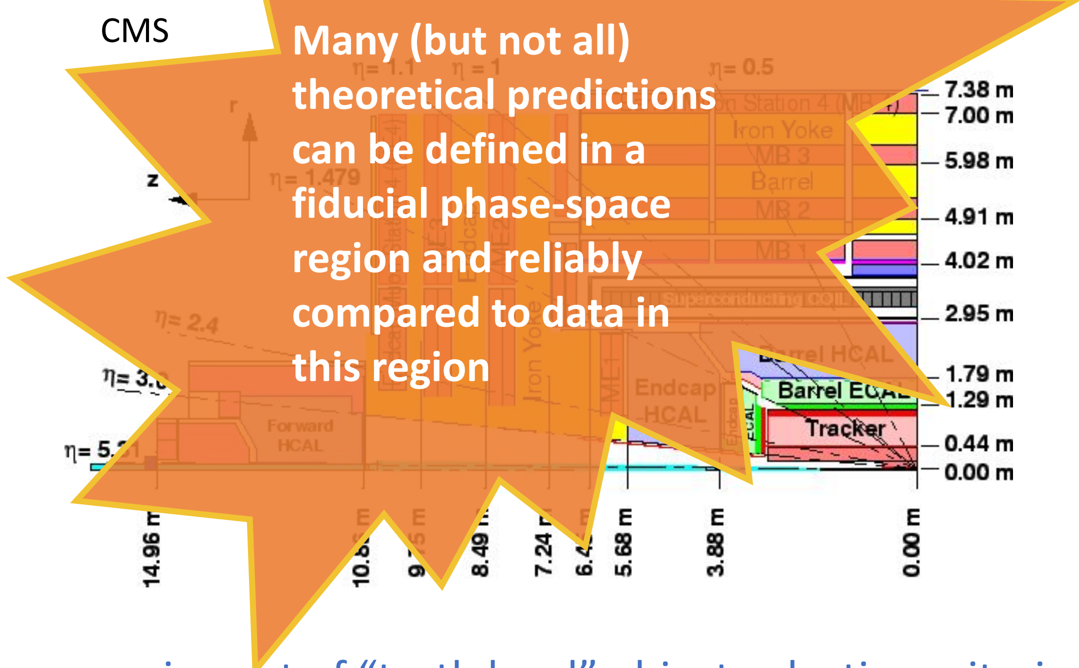
A fiducial phase-space is a set of "truth level" object-selection criteria that align well with the sensitivity of the real detector + reco:

e.g.: Select events with one (and only one) muon with p_T> 25 GeV, |η|<2.4 and p_T^{miss} > 30 GeV.

Fiducial phase-space

Regardless of detector efficiencies and resolutions, there are uninstrumented kinematic regions that we don't measure *at all*:

We do not have 4π detectors, and can't reconstruct particles down to zero pT!



Many (but not all) theoretical predictions can be defined in a fiducial phase-space region and reliably compared to data in this region

A fiducial phase-space is a set of "truth level" object-selection criteria that align well with the sensitivity of the real detector + reco:

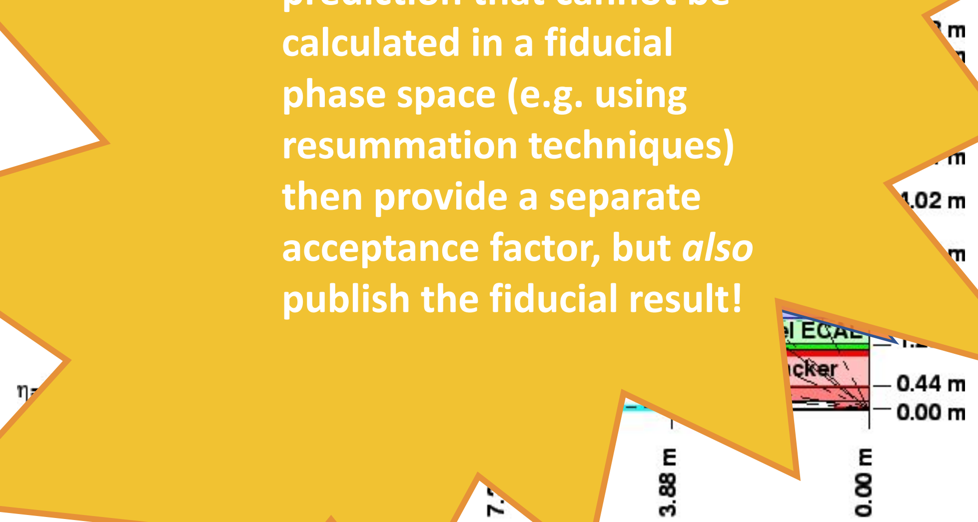
e.g.: Select events with one (and only one) muon with $p_T > 25$ GeV, $|\eta|<2.4$ and $p_T^{miss} > 30$ GeV.

Fiducial phase-space

Regardless of detector efficiencies and ... tic ... are uni ... trumented kinematic regions that we ... ’t measu ...

We do not have 4π detect ... !

If you have a theory prediction that cannot be calculated in a fiducial phase space (e.g. using resummation techniques) then provide a separate acceptance factor, but *also* publish the fiducial result!



A fiducial phase-space is a ... of “truth level” object-selection criteria that align well with the sensitivity of the ... eal detector + reco:

e.g.: Select events with one (and only one) muon with p_T> 25 GeV, |η|<2.4 and p_T^{miss} > 30 GeV.

What is a final-state particle?

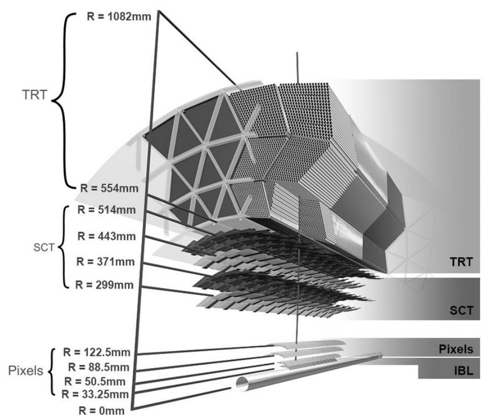
Remember…

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!

Common choice : $\tau_0 > 30$ ps, after hadronization.

Lots of useful advice and discussion here: https://cds.cern.ch/record/2022743

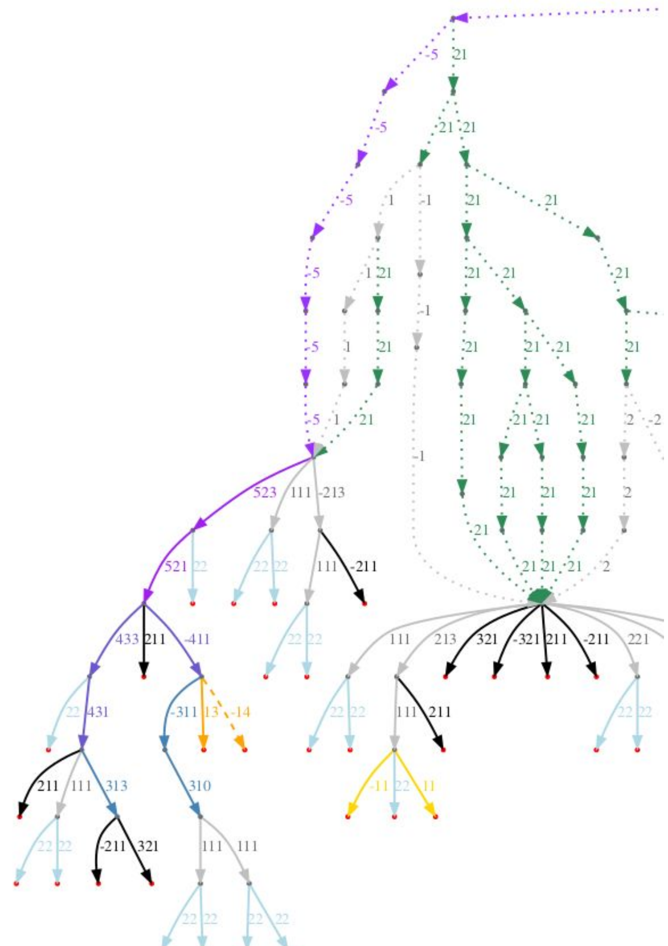


What is promptness / directness?

Not all final-state particles are equally significant:

- Hadrons are ten-a-penny! Even c and b hadrons can be produced fairly easily from semi-soft g ⟶ bb splittings
- Similar for photons: a high-energy photon direct from the hard process is a significant EM-interaction event, but *lots* of low-energy ones from $\pi^0 \longrightarrow \gamma\,\gamma$
- It’s harder to make electrons and muons come out of hadron collisions! Implies an electroweak process internally ⇒ weak coupling. Can still happen in semi-leptonic (esp. heavy) hadron decays, with neutrinos.
- Leptonic taus are indistinguishable from e/μ+ν

- Define *directness* “backwards”: recursively eliminate anything from a hadron. Only leptons and photons can be unambiguously direct



Often see “prompt” used to mean direct… but misleading since not a timing or displacement thing, unlike b-physics usage

What is a final-state electron/muon?

- Electrons/muons from hadron decays are typically removed in the data analysis by isolation cuts / fake removal:

✔ Can define direct leptons via the not-from-hadron-decays rule and only consider these: more robust and model-independent than asking that the lepton comes from a certain propagator in the hard process

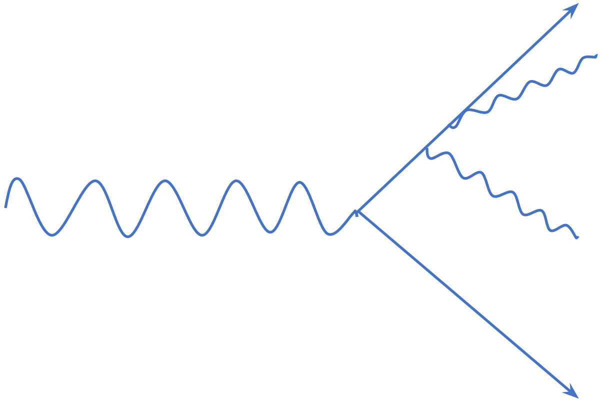
✔ **Well-defined in Rivet**, but you may need to also implement it in your experiment's software

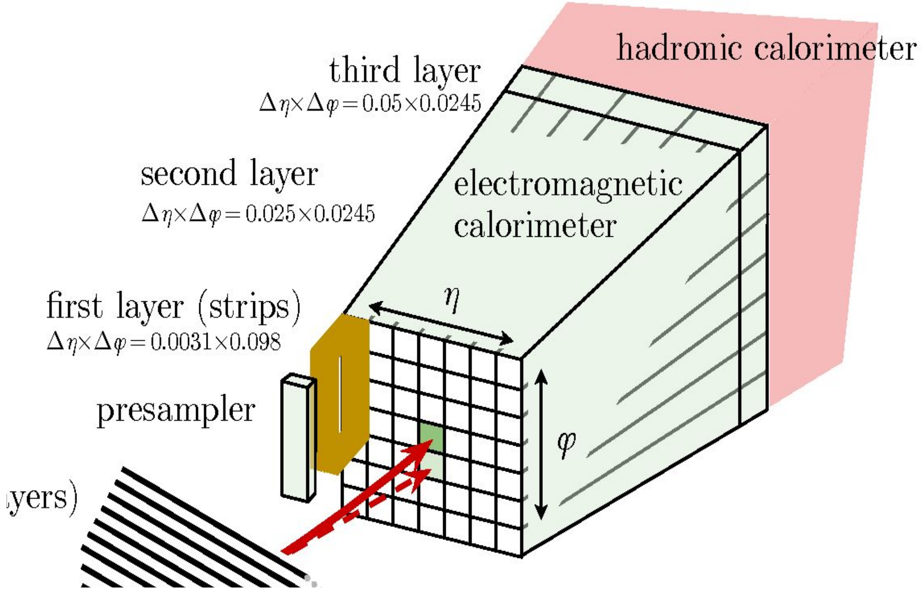
- We don't usually define particle-level isolation, but rather correct for inefficiencies of these requirements
  - It might be worth reconsidering this in specific analyses where proximity to jets has a large effect on results

What is a final-state electron/muon?

- Electrons and muons emit FSR photon radiation
  - (and lots of it, especially in the collinear limit, especially for electrons)

- For muons we measure the charged particle track, photon energy is not included
- For electrons we cluster calorimeter cells, and most collinear radiation will be included in the energy measurement

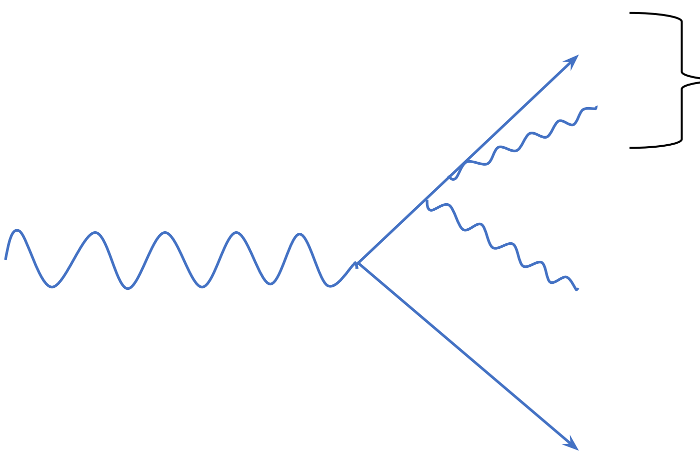




What is a final-state electron/muon?

- Electrons and muons emit FSR photon radiation (and lots of it, especially in the collinear limit, especially for electrons)

- We can define lepton momenta as:
  - **Born leptons** – as if FSR never happened ⇒ not what we measure
  - **Bare leptons** – after all FSR ⇒ closest to muon measurement
  - **Dressed leptons** – with the momenta of close-by photons “clustered” into the lepton momenta ⇒ closest to electron measurement



dressed : typically a $\Delta R < 0.1$ cone is used, but a jet algorithm may be better

What is a final-state electron/muon?

- Electron and muon final states can be very different for bare leptons, but much closer for Born and dressed leptons: see the Rivet tutorial

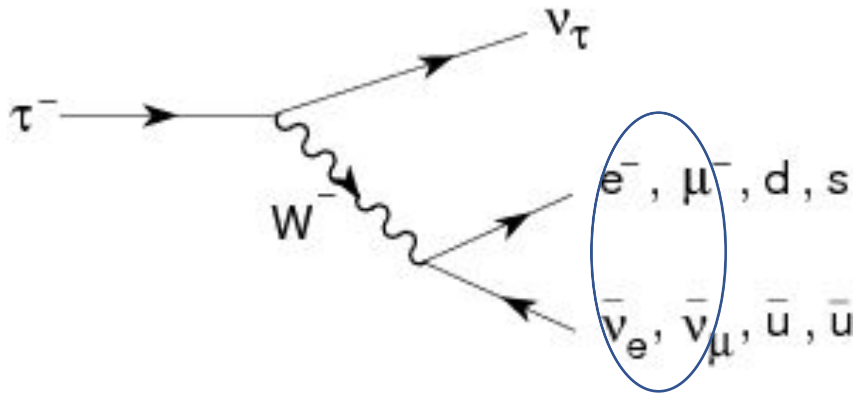
- It is often argued that dressed should be used for both to allow for *easy combination of final states.* Also bare versus dressed is much closer for muons than bare versus dressed for electrons

- Similarly, fiducial phase space cuts often harmonized for the two, requiring a small extrapolation in phase space for one

- But electrons ≠ muons
- We may want to retain sensitivity to differences
    - Especially cf. lepton (non-)universality anomalies!
- It may be better to measure both, publish correlations between uncertainties, and make choices that are best for each individual channel…

What is a final-state tau?

Recall: unstable ($c\tau_0$=0.1 mm, decays after ~1 mm at 20 GeV)

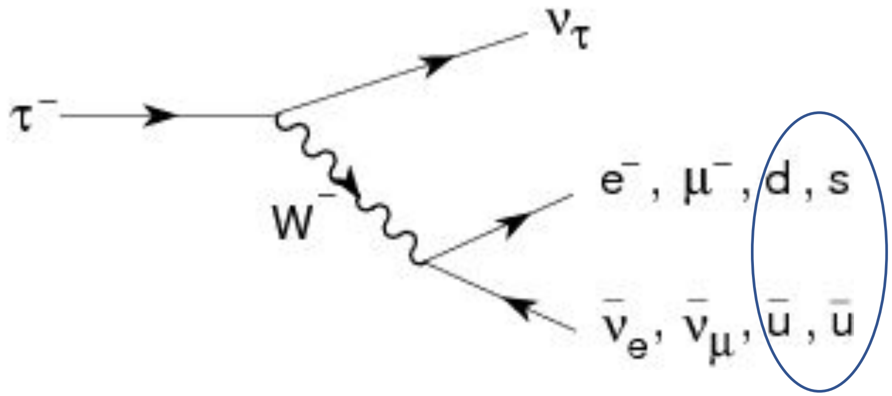


Leptonic decays

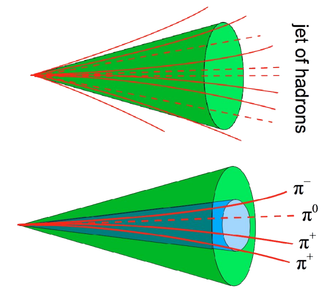
- The final state particles are electrons/muons and neutrinos
- Define fiducial phase-space with those (but we need to be careful to check lepton efficiencies as, e.g. impact-parameter cuts can be less efficient for leptons from taus)

What is a final-state tau?

Recall: unstable ($c\tau_0$=0.1 mm, decays after ~1 mm at 20 GeV)



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

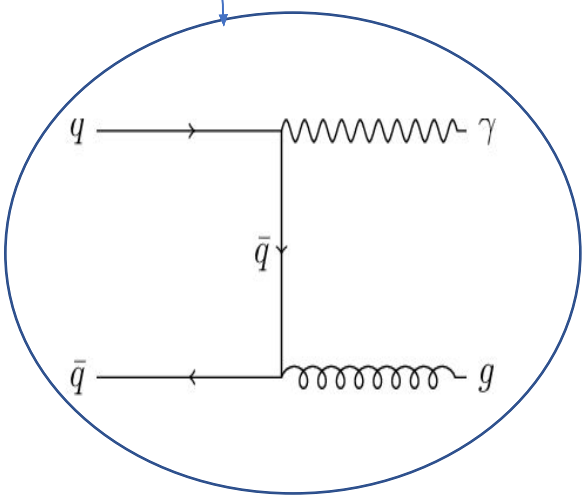


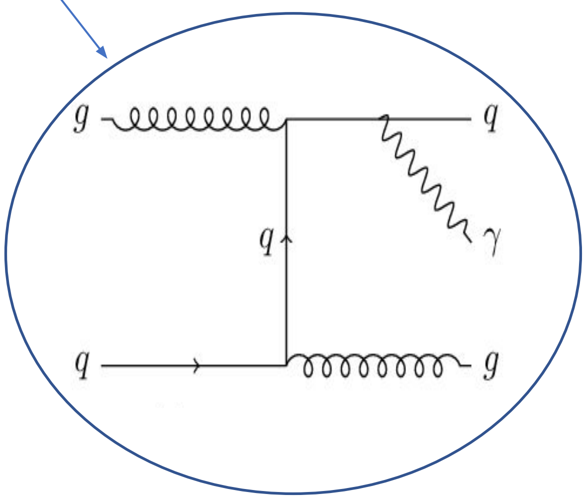
Hadronic decays

- Final-state particles are hadrons (⟶ jets) and neutrinos
- Complicated by the large number of hadrons *not* from taus. Experimental cuts reject backgrounds based on features of the jets: hard to replicate at particle-level
- A compromise might be best: require a hadron in the jet to have come from a direct tau (this is final-state based *in principle*).
- Not much experience! More detailed studies would be interesting

What is a final-state photon?

- Analyses usually measure direct/prompt, isolated photons
- Recall: direct or prompt mean “not from a hadron decay”
- But photons can be further divided into those from the hard scatter and those from parton fragmentation & non-perturbative hadronisation





A particle-level isolation criteria is necessary to replicate the isolation applied at reco-level. Maybe jet-based. See also problems at fixed-order, cf. Frixione and soft-drop isolation

Note in principle this could also be done for prompt leptons, but it is much less important … *why?*

What is a final-state neutrino/invisible?

Invisible in the detector and existence inferred by p_T^{miss}

Recall:

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

- Sometimes the momenta of (prompt?) invisible* particles are summed
- An alternative is to take – the sum of all the visible particles within detector acceptance, which is closer to what we measure but can be a bit complicated. E.g. what p_T of hadrons are we actually sensitive to? And what about reco calibrations? (More on this later)

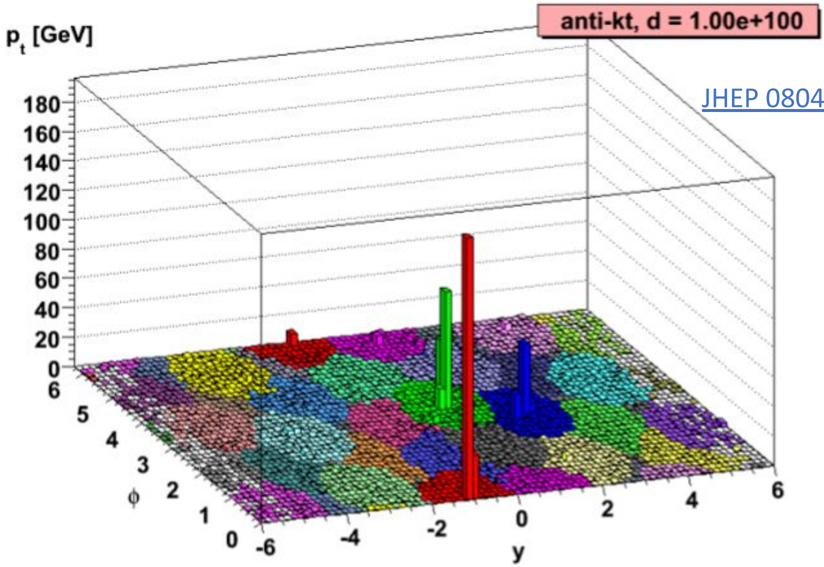
* neutrinos are indistinguishable from BSM invisible particles, and multiple invisibles are indistinguishable from single invisibles

What is a final-state parton?

- Partons radiate more partons, recursively. They all eventually hadronize
- Run a jet algorithm on the final-state particles
  - Form a list of particles (this would be clusters / tracks at reco-level)
  - Merge the smallest pair according to a “distance” parameter
  - Iterate (until a stopping condition, cf. the “beam distance”)
- Algorithms assign each hadron to a jet. The energy/momentum of the jet represents the energy/momentum of the parton from the hard scatter
- Think carefully about what is included as inputs: Muons? Neutrinos?

Note: In principle an electron will always form a jet experimentally.

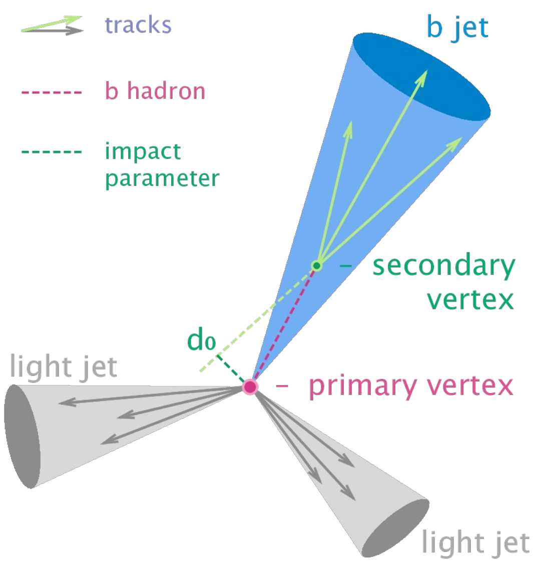
We remove these jets using overlap removal at both reco- and truth-level (e.g. remove any jets with △R < 0.4 from a prompt electron)



JHEP 0804:063,2008

What is a final-state *b*-quark/jet?

- Recall decay length for a 20 GeV *b*-hadrons ~2 mm, they are therefore unstable and not included as final-state particles
- But we select them experimentally by making displaced-vertex selections



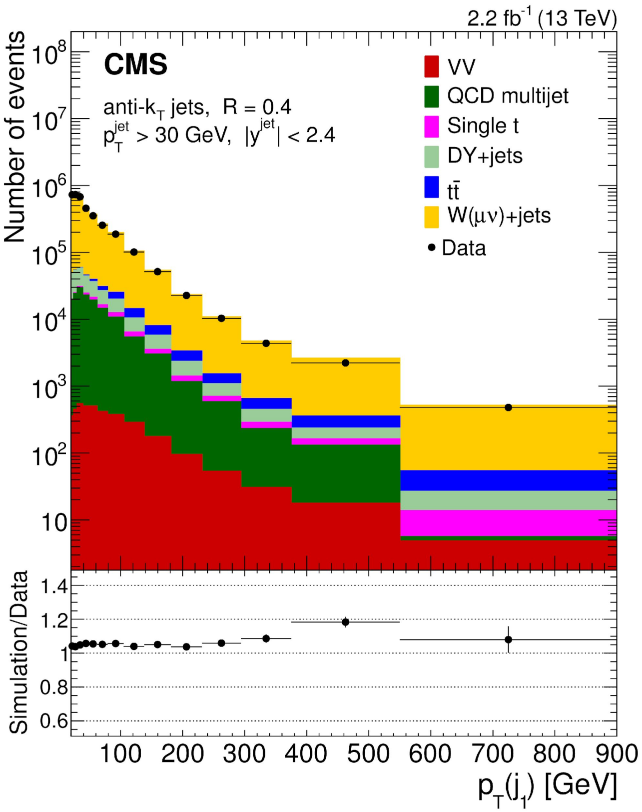
- Common "compromise" is to associate the *non-final state b*-hadrons to jets
  - More "in-principle final-state": it *could* have been reconstructed
- If a jet contains a *b*-hadron it is considered a particle-level *b*-jet
  - Maybe with a p_T cut on the *b*-hadron, and careful definition of "contains"

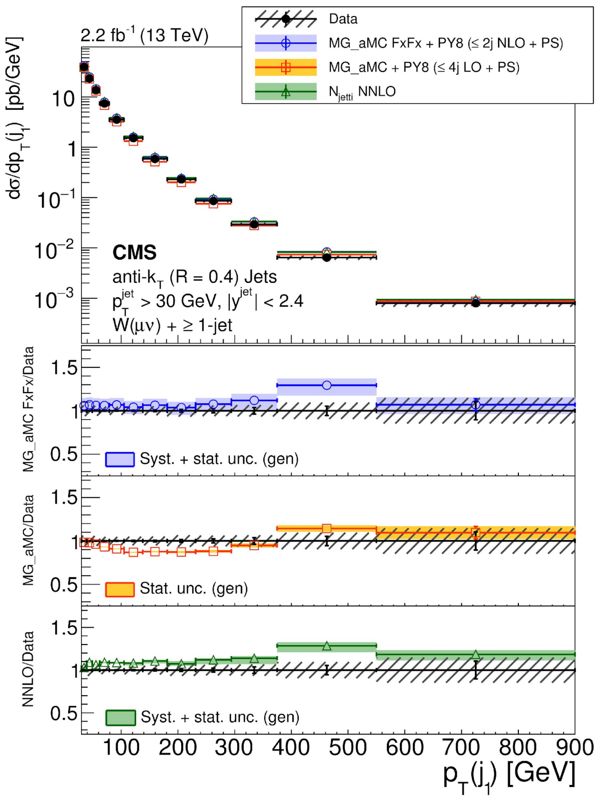
Examples: CMS *W*+jets

Phys. Rev. D 96 (2017) 072005
Rivet: CMS_2017_I1610623

Fiducial phase-space: (Follow data-analysis cuts closely)

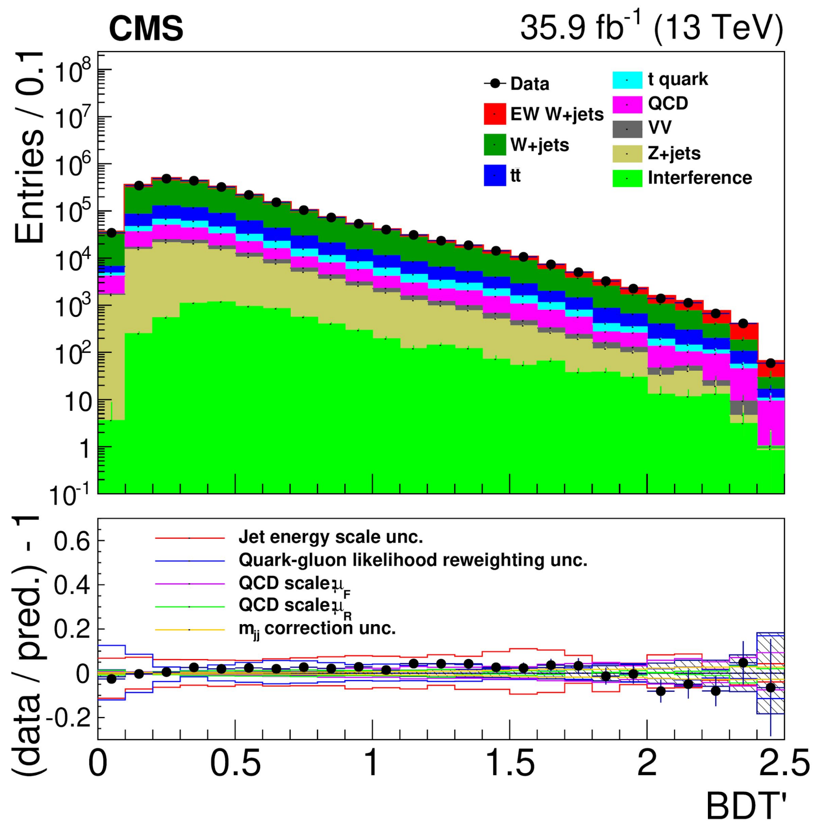
- One dressed ($\Delta R < 0.1$) prompt muon with $p_T > 25$ GeV, $|\eta|<2.4$
- $m_T > 50$ GeV (using muon and prompt truth neutrino)
- Jets (exclude neutrinos and above muon): anti-k_T (R=0.4) with $p_T > 30$ GeV, $|y|<2.4$, and $\Delta R > 0.4$ from the muon





Examples: CMS EWK *W*+dijet

arXiv:1903.04040 (Submitted to EPJC)
No Rivet routine



Data analysis cuts:
jet 1 p_T > 50 GeV, jet 2 p_T > 30 GeV
m_{jj} > 200 GeV
p_T balance cut
p_T lepton > 20 GeV, $|\eta| < 2.4$

Fiducial phase-space:
lν + dijets final state
m_{jj} > 120 GeV and jet p_T > 25 GeV

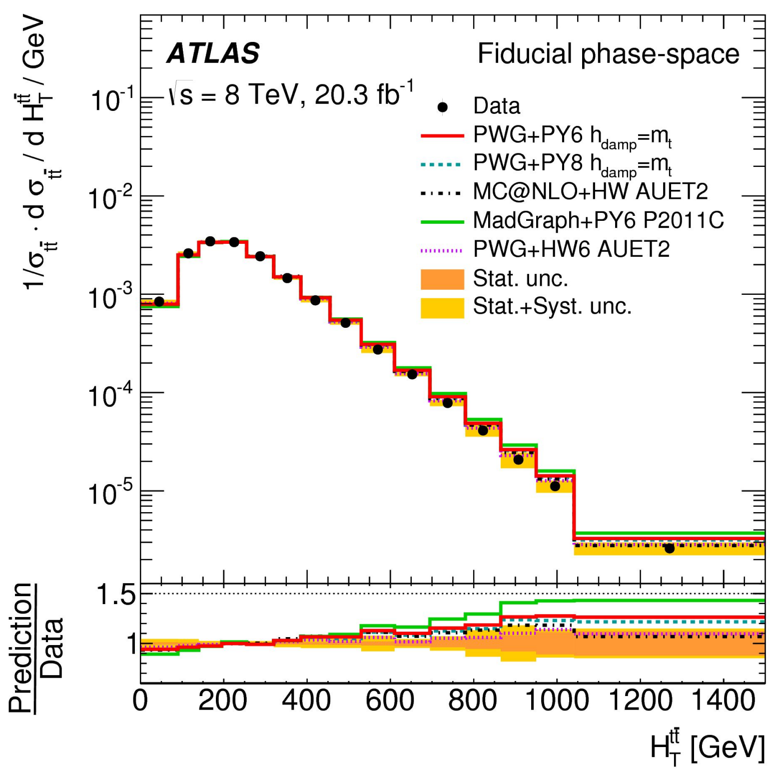
$\sigma_{EW}(Wjj) = 6.23 \pm 0.12\ (\text{stat}) \pm 0.61\ (\text{syst})$ pb

🤔 Why such a large extrapolation?

Examples: ATLAS top-quark analyses

[Eur. Phys. J. C76 (2016) 538](#)

Rivet: ATLAS_2015_I1404878



Eur. Phys. J. C 78 (2018) 487
DOI: 10.1140/epjc/s10052-018-5904-z
CERN-EP-2017-276
27th June 2018

Measurement of the inclusive and fiducial $t\bar{t}$ production cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector

The ATLAS Collaboration

The inclusive and fiducial $t\bar{t}$ production cross-sections are measured in the lepton+jets channel using 20.2 fb^{-1} of proton–proton collision data at a centre-of-mass energy of 8 TeV recorded with the ATLAS detector at the LHC. Major systematic uncertainties due to the modelling of the jet energy scale and b-tagging efficiency are constrained by separating selected events into three disjoint regions. In order to reduce systematic uncertainties in the most important background, the W+jets process is modelled using Z+jets events in a data-driven approach. The inclusive $t\bar{t}$ cross-section is measured with a precision of 5.7% to be $\sigma_{\text{inc}}(t\bar{t})$ = 248.3 ± 0.7 (stat.) ± 13.4 (syst.) ± 4.7 (lumi.) pb, assuming a top-quark mass of 172.5 GeV. The result is in agreement with the Standard Model prediction. The cross-section is also measured in a phase space close to that of the selected data. The fiducial cross-section is $\sigma_{\text{fid}}(t\bar{t})$ = 48.8 ± 0.1 (stat.) ± 2.0 (syst.) ± 0.9 (lumi.) pb with a precision of 4.5%.

ATLAS-CONF-2019-029

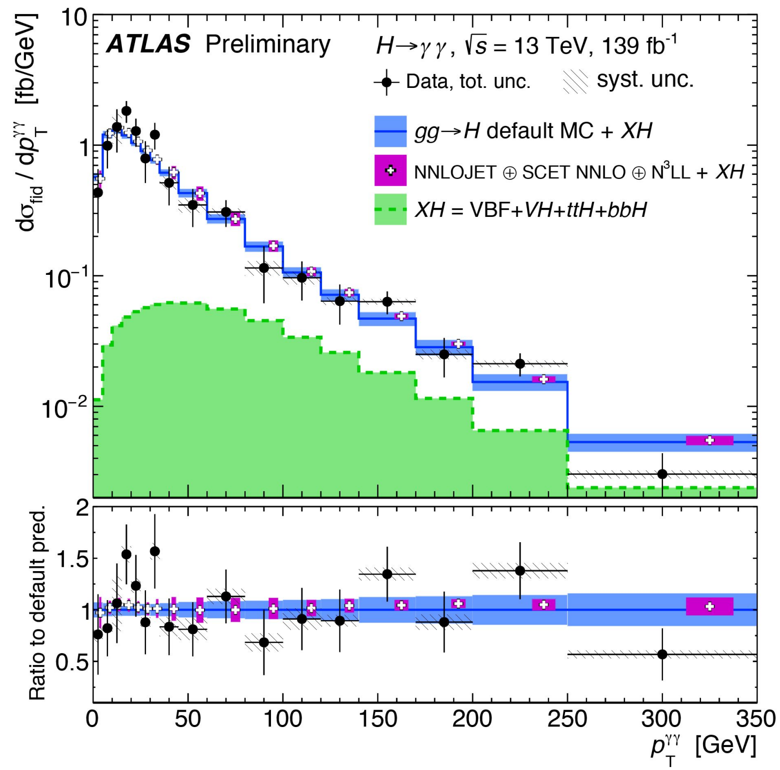
Examples: ATLAS Higgs → γγ

Table 2: Particle-level selections for the fiducial measurements. The photon isolation, $\sum p_{\mathrm{T}}^{i}/p_{\mathrm{T}}^{\gamma}$, is defined as the sum of the p_{T} of charged particles within $\Delta r < 0.2$ of the photon.

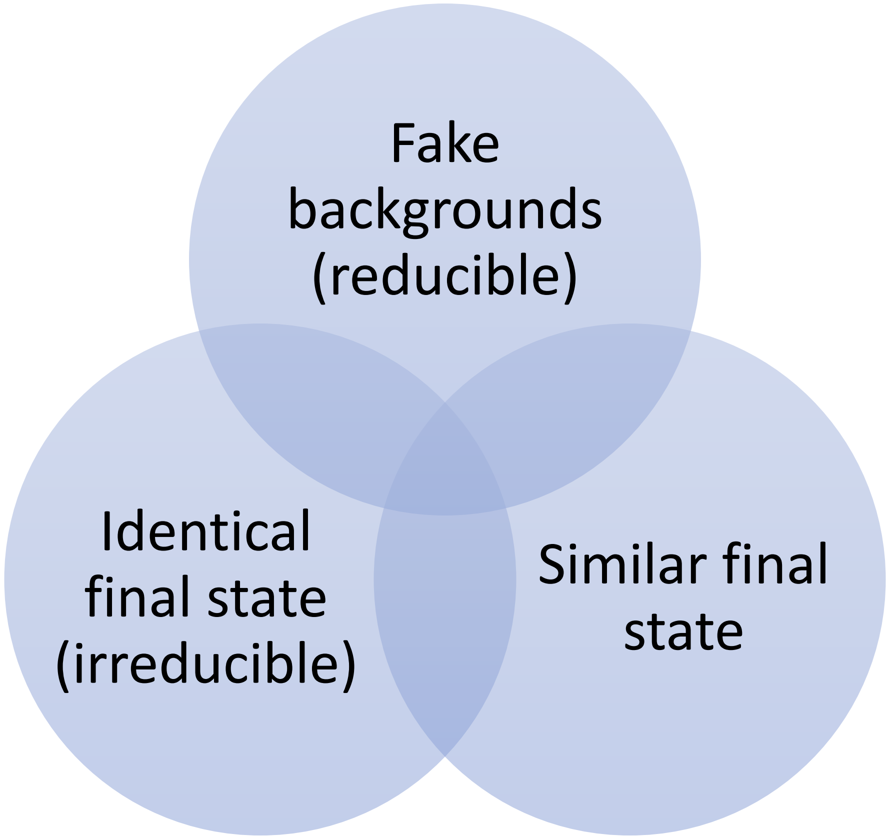
Objects	Fiducial definition
Photons	$\lvert\eta\rvert < 2.37$ (excluding $1.37 < \lvert\eta\rvert < 1.52$), $\sum p_{\mathrm{T}}^{i}/p_{\mathrm{T}}^{\gamma} < 0.05$
Jets	anti-k_t, $R = 0.4$, $p_{\mathrm{T}} > 30$ GeV, $\lvert y\rvert < 4.4$
Diphoton	$N_\gamma \geq 2$, $105\,\mathrm{GeV} < m_{\gamma\gamma} < 160\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\gamma_1}/m_{\gamma\gamma} > 0.35$, $p_{\mathrm{T}}^{\gamma_2}/m_{\gamma\gamma} > 0.25$

Fiducial Higgs!

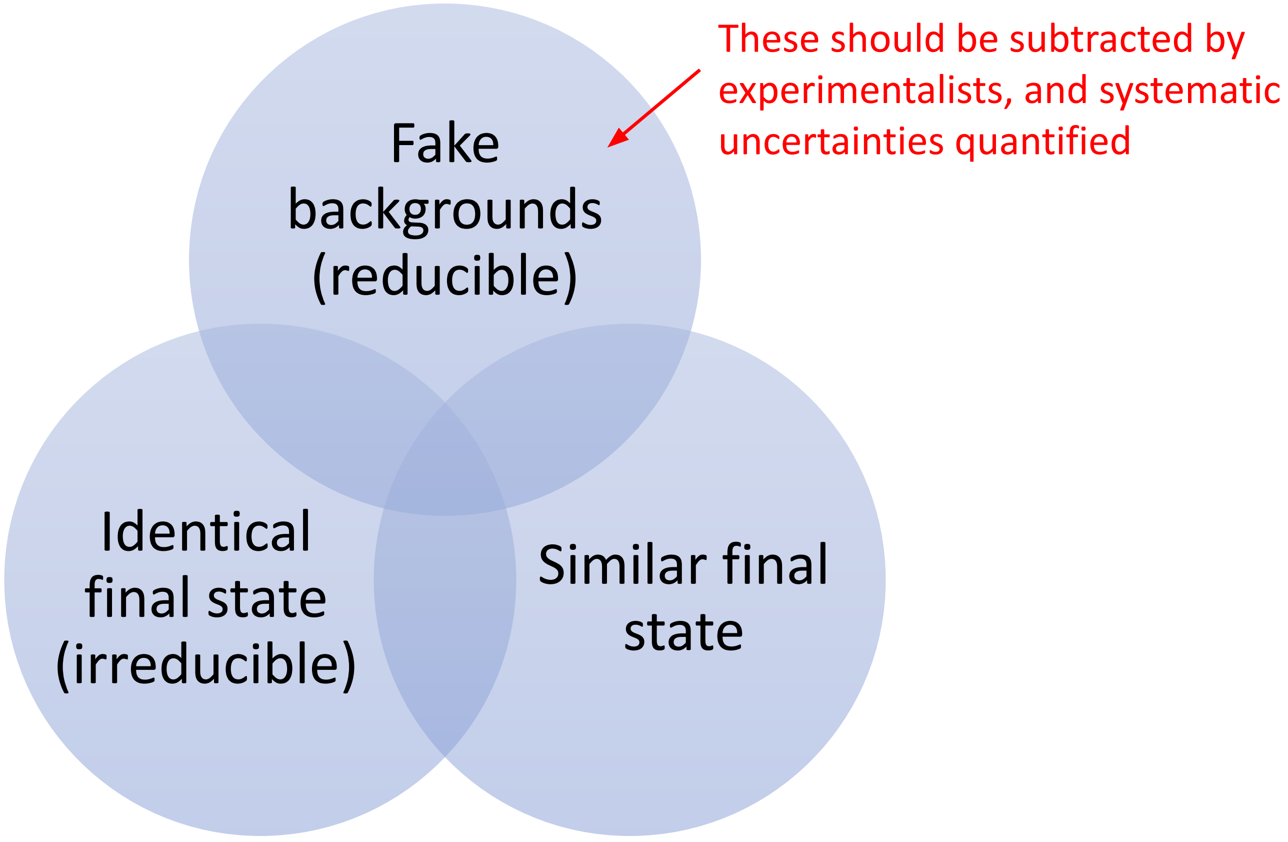
An important model-independent complement to other, more BSM-oriented approaches like STXS bins (per-production mode, summed over decays, often derived by unfolding from BDTs, etc.)



Background-subtraction… or not?



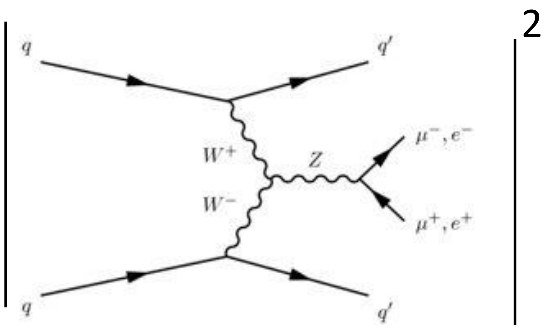
Background-subtraction… or not?

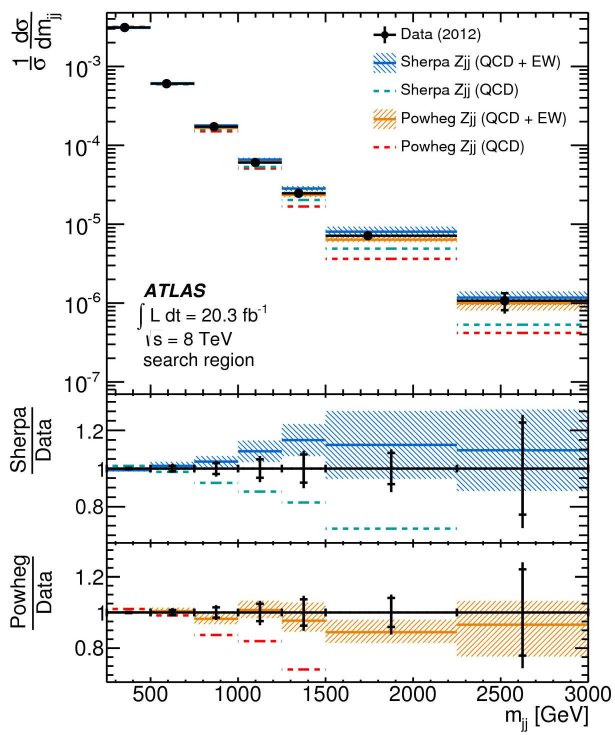


Identical final-states

- Quantum mechanics tells us that processes with identical final states will interfere and cannot be calculated separately
    - Usually. Defining “identical” can be subtle, cf. colour indices
- Sometimes this is a huge effect and separating out diagrams breaks gauge-invariance
- Other times the effect is quite small and attempts are made to isolate certain processes (e.g. diagram removal/subtraction)

Identical final-states example: $\ell^+ \ell^-$ VBF

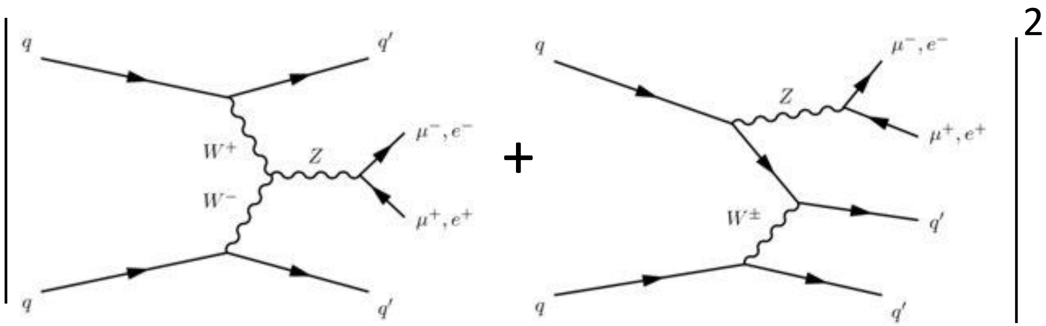


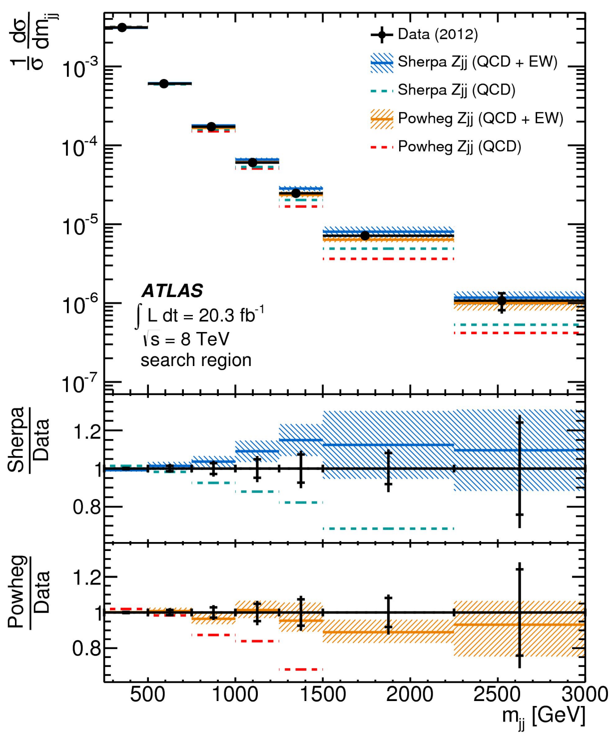


JHEP04(2014)031

Rivet: ATLAS_2014_I1279489

Identical final-states example: … is $\ell^+ \ell^-$ + dijet



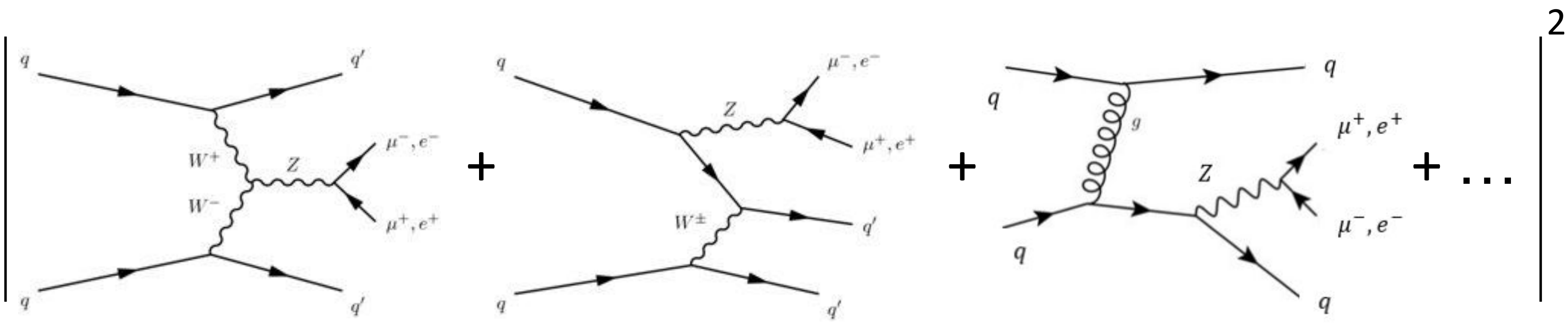


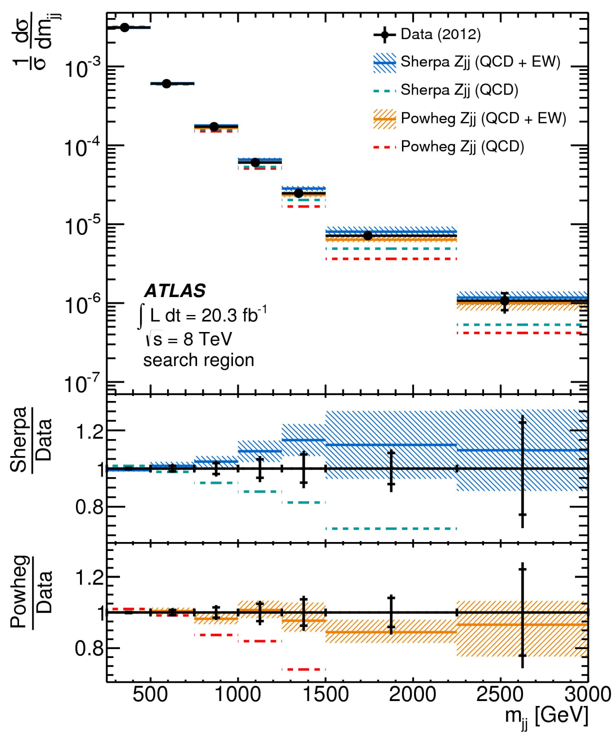
JHEP04(2014)031

Rivet: ATLAS_2014_I1279489

"The VBF process cannot be isolated due to a large destructive interference with the electroweak *Z*-boson bremsstrahlung process."

Identical final-states example: … etc.!



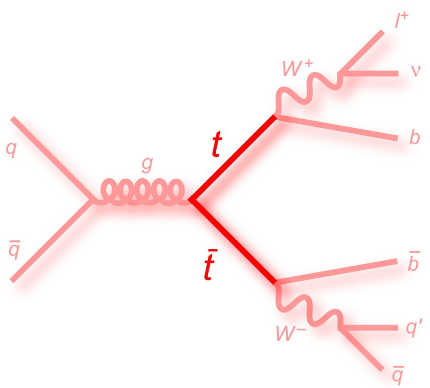


JHEP04(2014)031

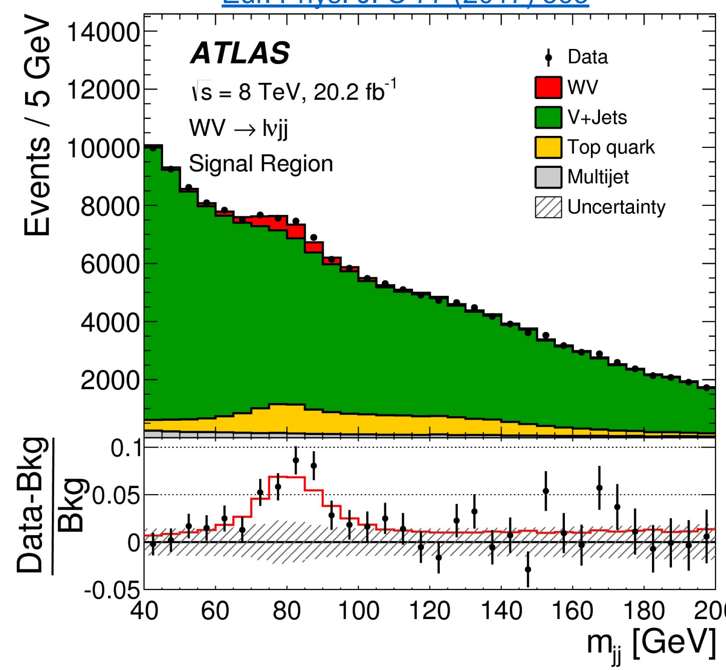
Rivet: ATLAS_2014_I1279489

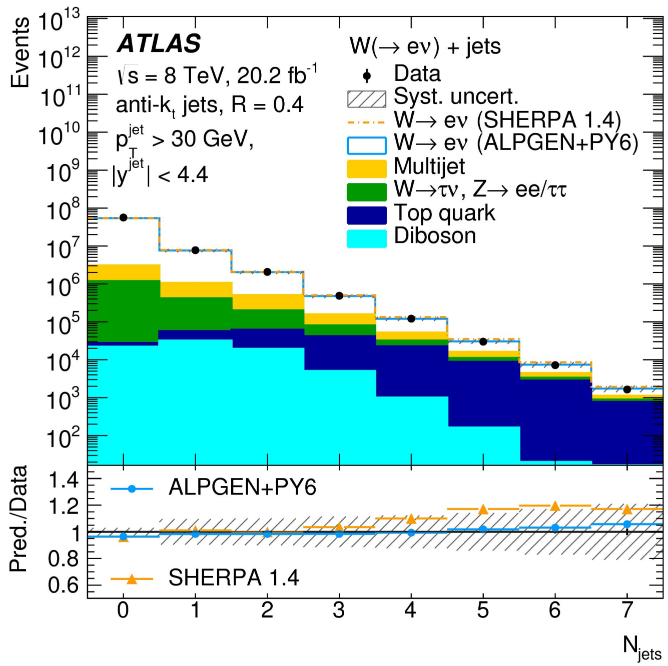
It's usually good to also include an inclusive measurement with no assumptions or subtractions made

And ttbar is *WW*+jets is *W*+jets…



[Eur. Phys. J. C 77 (2017) 563]

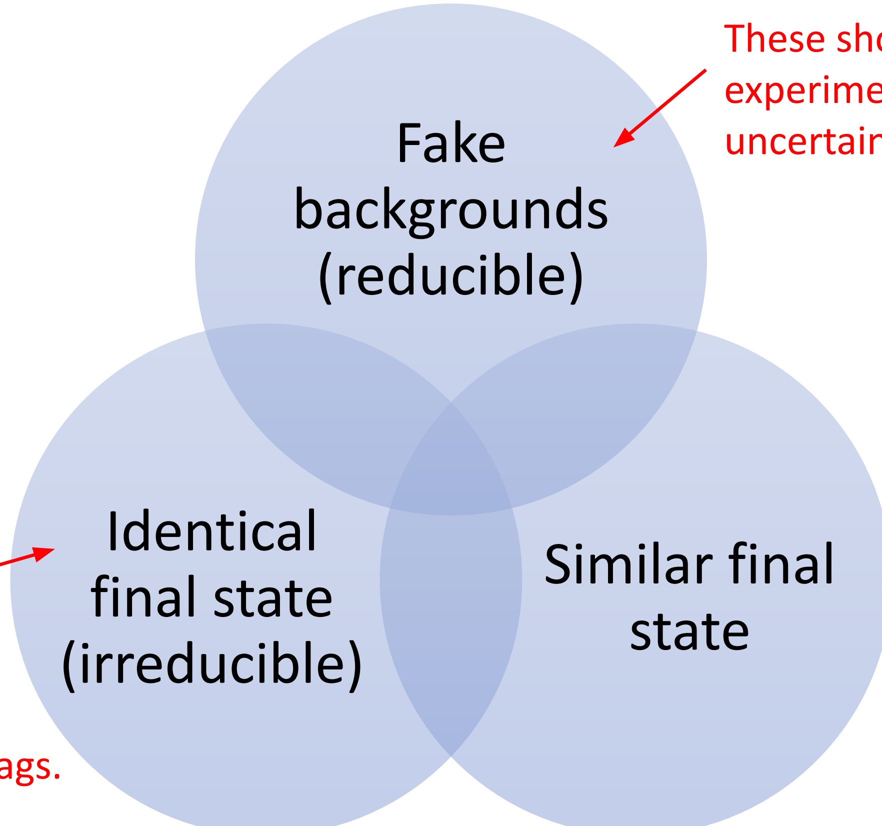




[JHEP 05 (2018) 077]

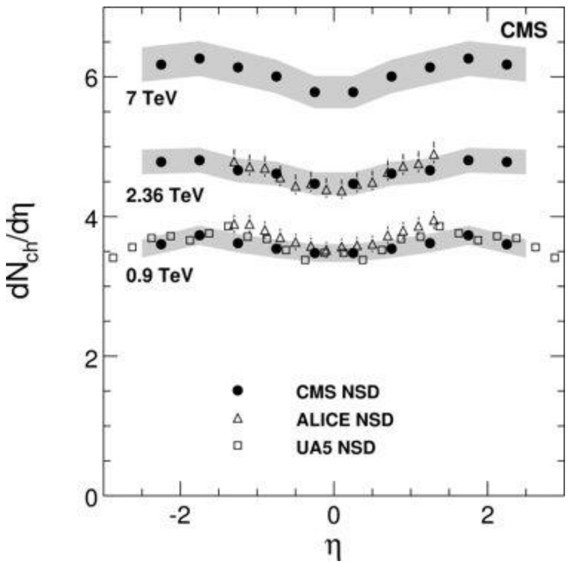
So, background-subtraction… or not??

These should be subtracted by experimentalists and systematic uncertainties quantified



In general it is dangerous to isolate certain amplitudes/diags. In some cases it is a reasonable approximation but treat with caution and try to measure an inclusive observable too!

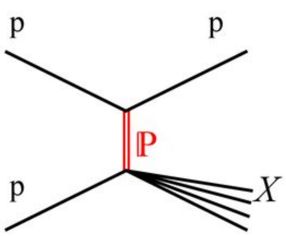
Similar final-states: soft-QCD



[Phys. Rev. Lett. 105 (2010) 022002](#)

Rivet: CMS_2010_S8656010

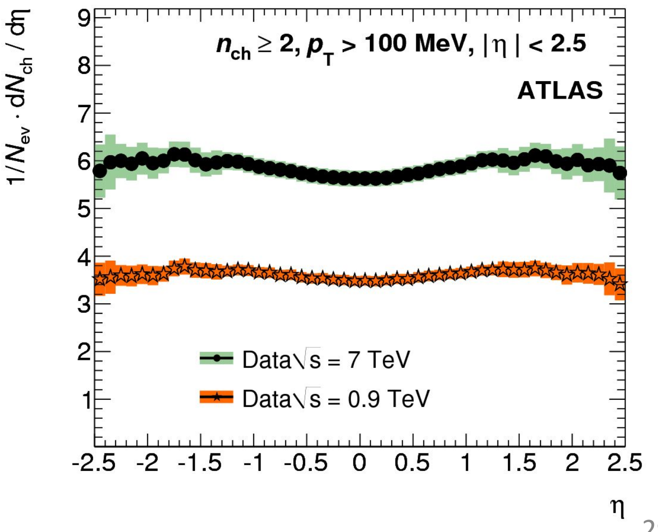
Single-diffractive subtracted!!!



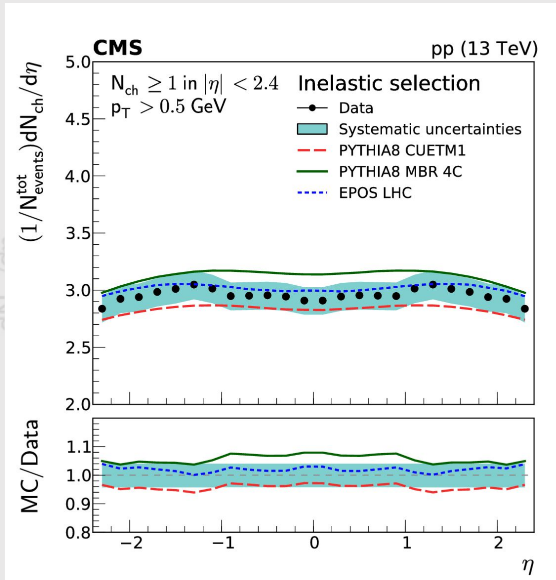
[New J. Phys. 13 (2011) 053033](#)

Rivet: ATLAS_2010_S8918562

Final-state particle definition



Similar final-states: soft-QCD

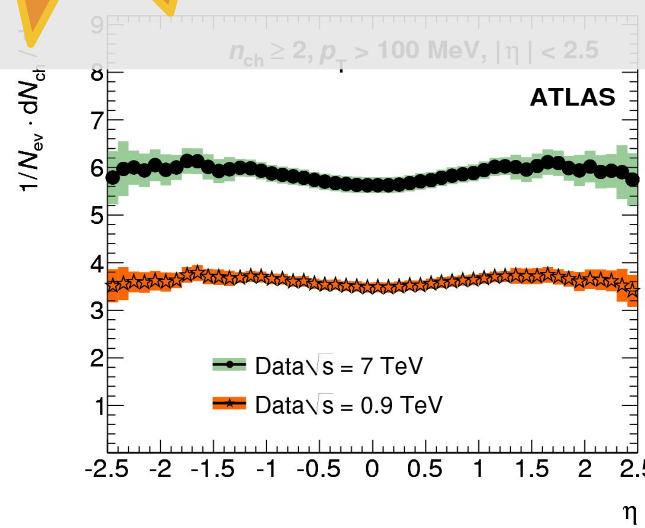


Eur. Phys. J. C 78 (2018) 697
Rivet: CMS_2018_I1680318

Final-state particle definition

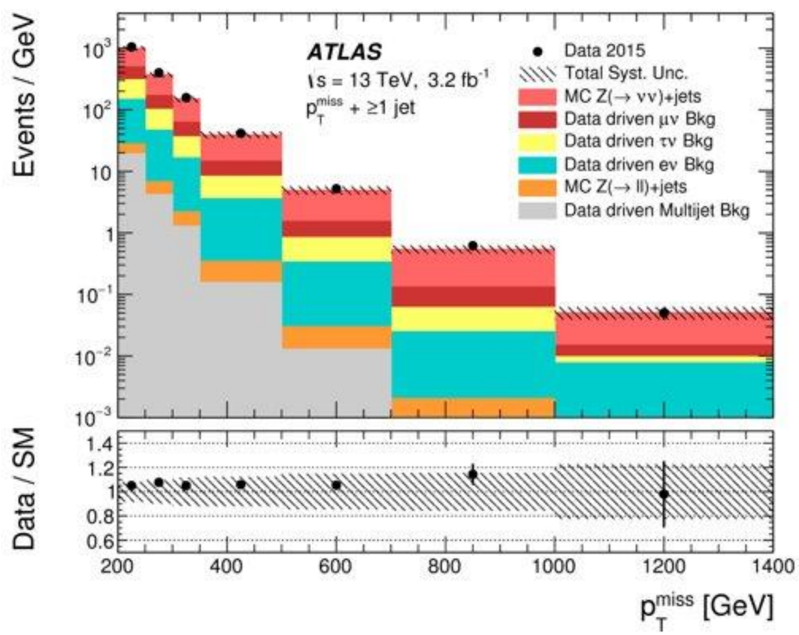
New J. Phys. 13 (2011) 053033
Rivet: ATLAS_2010_S8918562

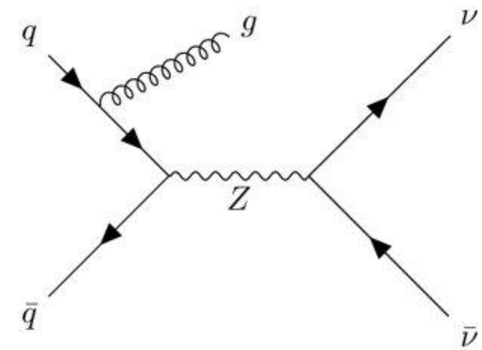
Final-state particle definition



Similar final-states: $p_{T,miss}$ + jets

Eur. Phys. J. C 77 (2017) 765
Rivet: ATLAS_2017_I1609448





Fiducial phase-space:
p_T^{miss} + jet(s)
No charged leptons with $|\eta|< 2.5$, $p_T > 7$ GeV
+ others

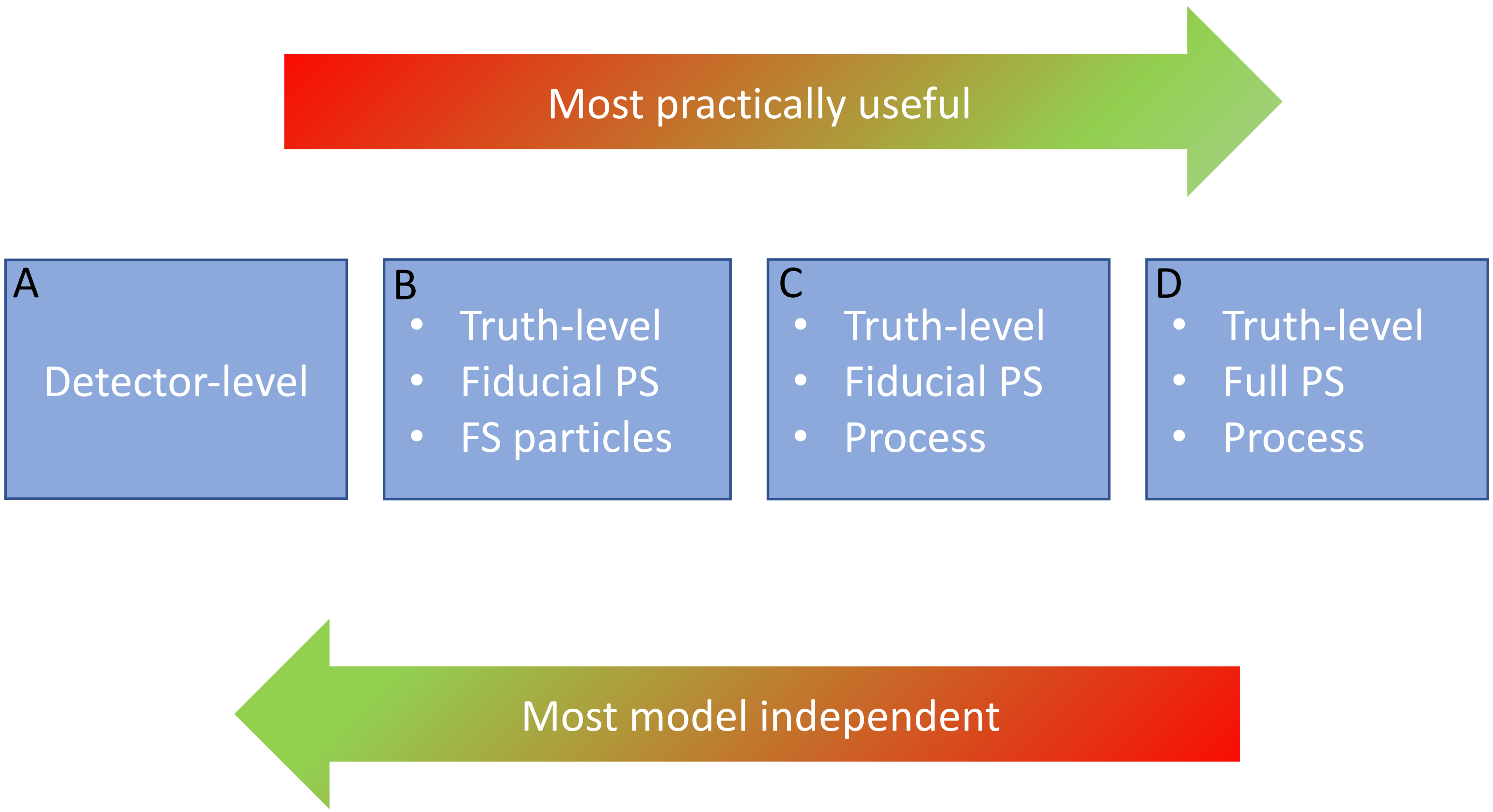
W→lν with "out of acceptance" leptons contribute ~ the same as Z→νν!

- In this paper: background determined using control regions+MC, and subtracted
- Perhaps these *W*'s should be included as part of the "signal" definition?
  - This leaves the data uncontaminated and as close to "what we see" as possible.
  - Removes dependence on control regions and MC extrapolation between regions
- But be careful of fiducial phase-space definitions: e.g. out-of-acceptance muons should be included as *invisible* in a particle-level p_T^{miss} definition!

What is the perfect measurement?

What is the ~~perfect~~ optimal measurement?

What is the ~~perfect~~ optimal measurement?



optimal

What is the ~~perfect~~ measurement?

Most practically useful →

A

Detector-level

B

- Truth-level
- Fiducial PS
- FS particles

C

- Truth-level
- Fiducial PS
- Process

D

- Truth-level
- Full PS
- Process

My favourite (B)

Most common (C)

← Most model independent

BSM searches & detector-correction

- Typically done at reco-level, searches set limits on parameters in a given model by comparing to *reco-level* MC predictions
- But data in a given analysis can be sensitive to *many* BSM theories.
  How to re-interpret the measurements?
  - Many people working on how to reinterpret reco-level results, e.g. by using fast detector simulation (can be interfaced with Rivet)
  - Another option is to correct for detector effects and allow comparisons with "truth-level" predictions. Some sensitivity may be lost due to binning but much easier to reinterpret

Distinction between BSM search and SM measurement becomes blurred:

- We measure the data in certain final-states and compare to the best SM predictions.
- We should do this more in regions particularly sensitive to new physics
- Important to stick to the "measuring a final-state" philosophy

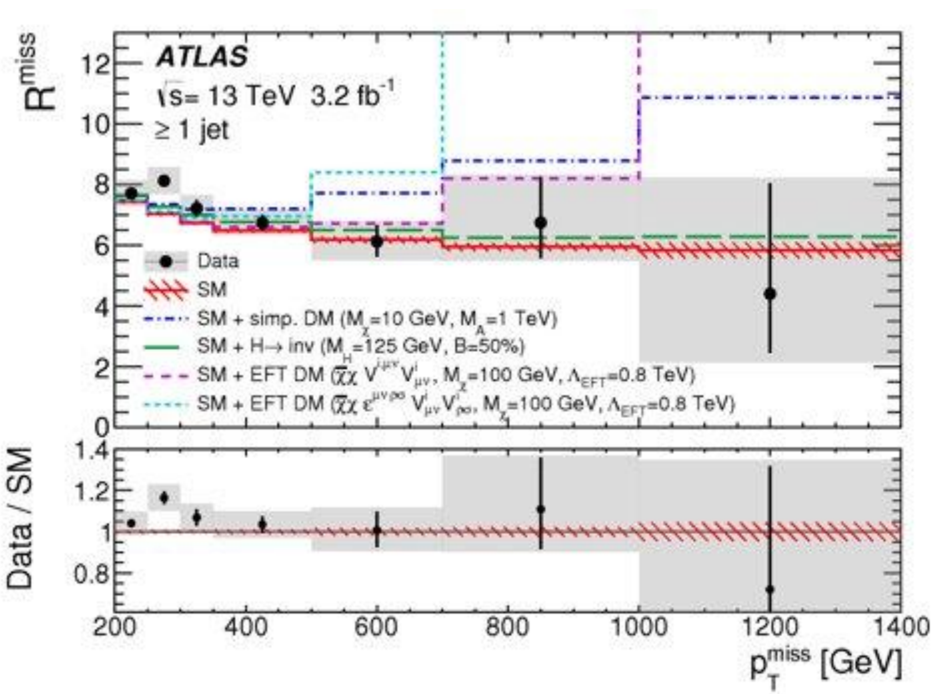
See Contur https://contur.hepforge.org and its tutorial: arXiv:2102.04377

$\Rightarrow$ idea is to use all analyses in Rivet to constrain BSM parameter spaces

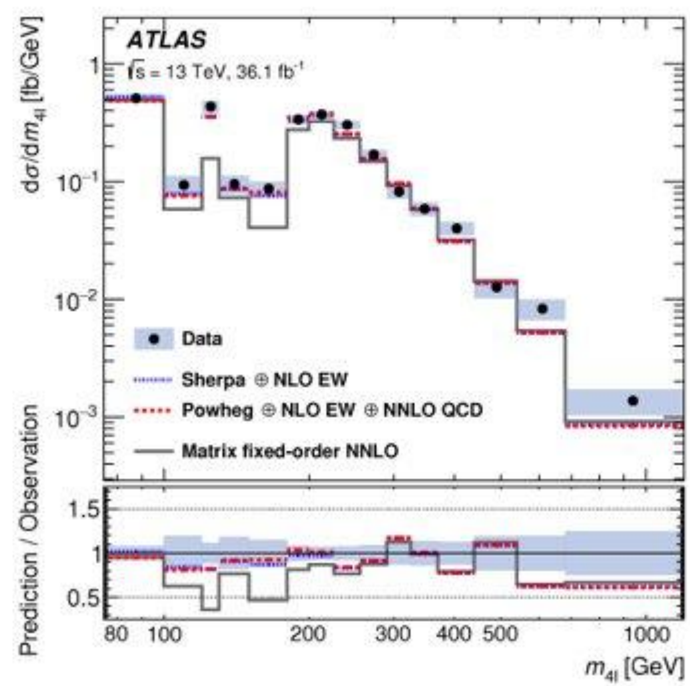
Sometimes surprises occur and a certain model pops up in multiple final states we haven't thought of yet!

BSM searches & detector-correction

Unfolding model-independent observables in BSM-search phase-space is a pretty new, and very exciting thing to do at the LHC!



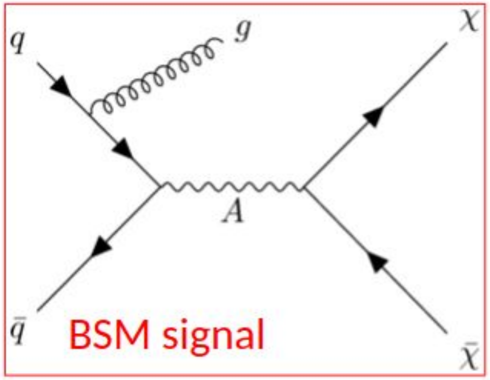
Eur. Phys. J. C 77 (2017) 765

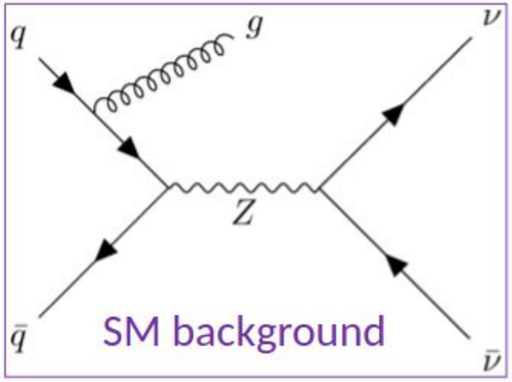


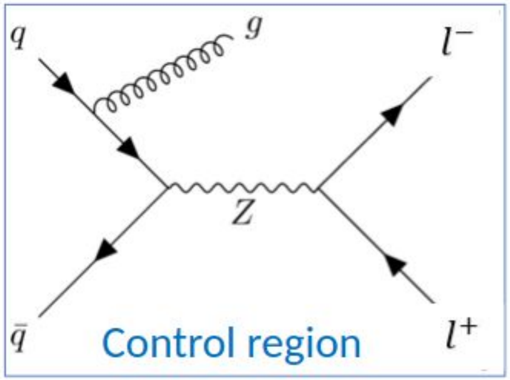
JHEP 04 (2019) 048

BSM searches & backgrounds

- Often backgrounds to BSM searches are predicted using constraints from “control regions”. These can be
  - similar final states, or
  - the same final-state with different kinematic cuts.
- This can be very useful, especially when modelling is bad, and can reduce systematic uncertainties a lot.







But it can limit re-interpretation: what if another BSM theory leads to final-state particles in the control region too?

“Everyone’s signal is someone else’s background”!

BSM searches: backgrounds & unfolding

A possible solution:

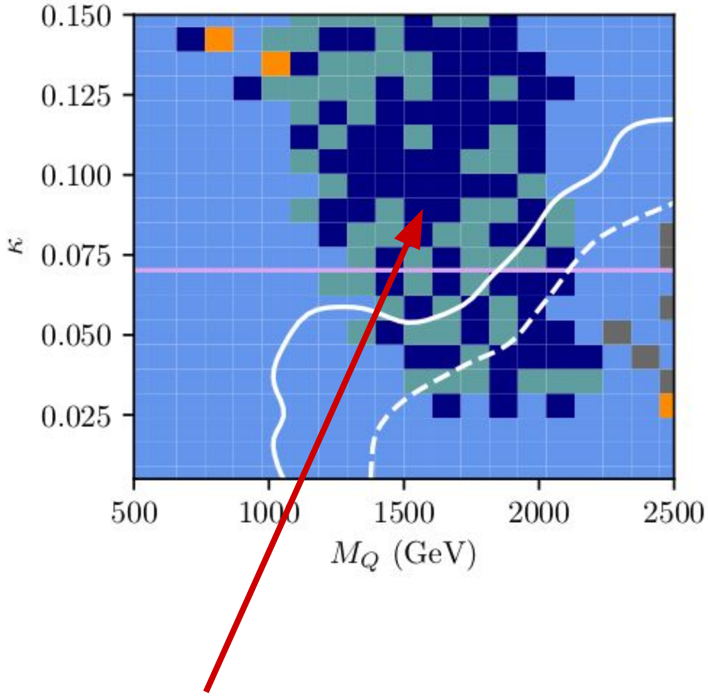
- Unfold and publish the signal region and the control region with correlation information
- Control region constraints can then be made for models that allow it but not for others

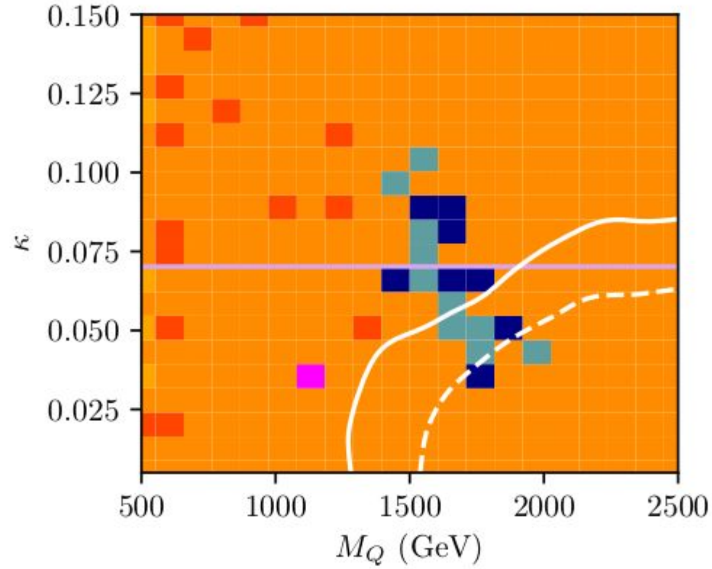
Sometimes, maybe we should just say what we see…

These are all (weirdly) new ideas.
There's lots of room for studies and analyses — get involved!

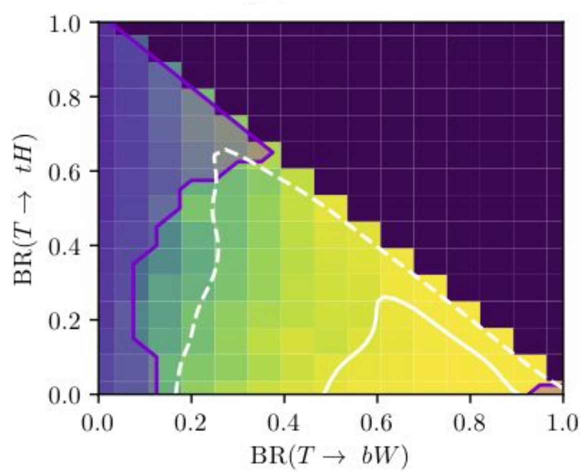
BSM searches & control-region unfolding

SciPost Phys. 9 (2020) 5, 069
[arxiv:2006.07172](https://arxiv.org/abs/2006.07172)





VLQ reinterpretation: unexpected "resonant" injection from a *Wjj* unfolded control-region in Rivet!

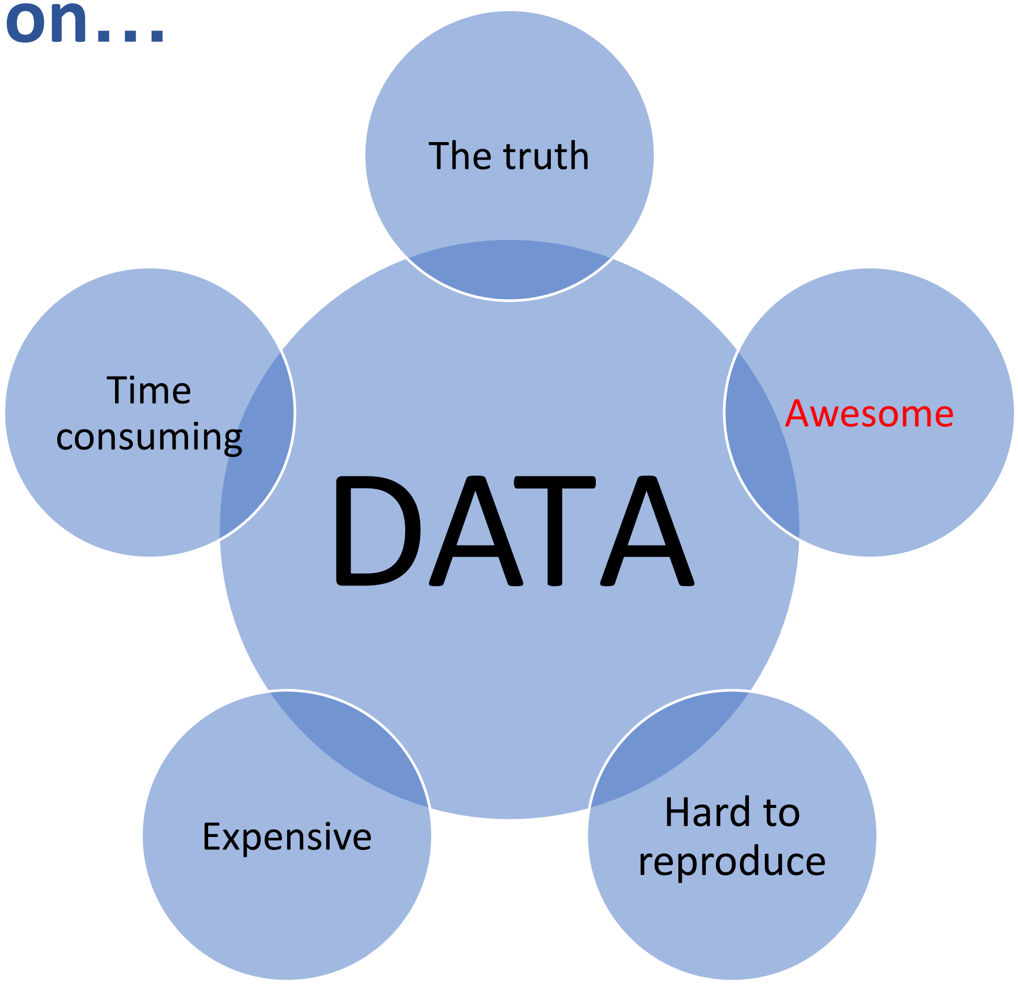


Another *search control-region* was also important in adding sensitivity!

Better experimental analysis: a summary

- ✔ Correct (carefully) for detector effects (maybe even for BSM searches)
- ✔ Measure your fiducial phase-space
- ✔ Think carefully about subtracting “backgrounds”
- ✔ Keep the data as clean and model-independent as possible

In conclusion…



… data is still awesome: look after it!