Physics objects for top physics in ATLAS

Richard Hawkings (CERN)

Top2016 conference, Olomouc, Czech Republic, 19/9/2016

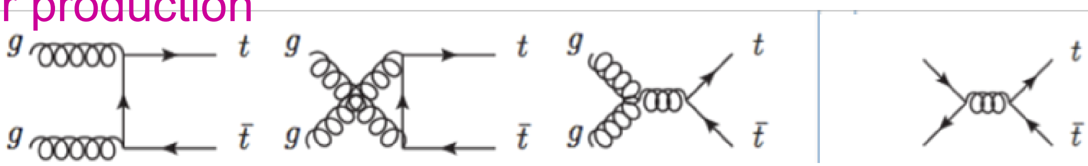
- The ingredients for top physics measurements
  - Data/detector aspects – MC generators / simulation covered elsewhere
- Top decays and associated objects
- Data samples for top physics
- The objects:
  - Electrons, muons, taus, jets, b-tagging, E_T^{miss}
- Summary

- All plots / results shown taken from ATLAS public web:
  - https://twiki.cern.ch/twiki/bin/view/AtlasPublic
  - In particular ‘combined performance’ group pages (see References at end)

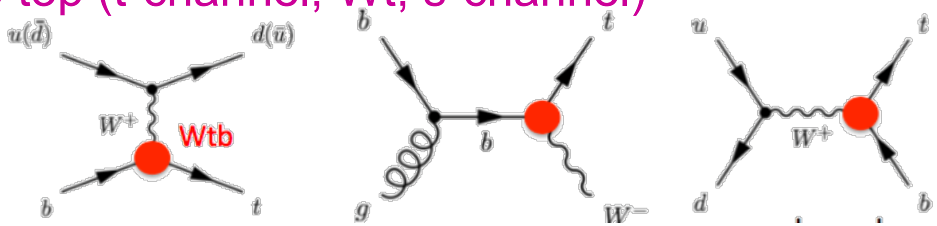
Anatomy of top events

- Tops produced in pairs (tT) or singly

Pair production

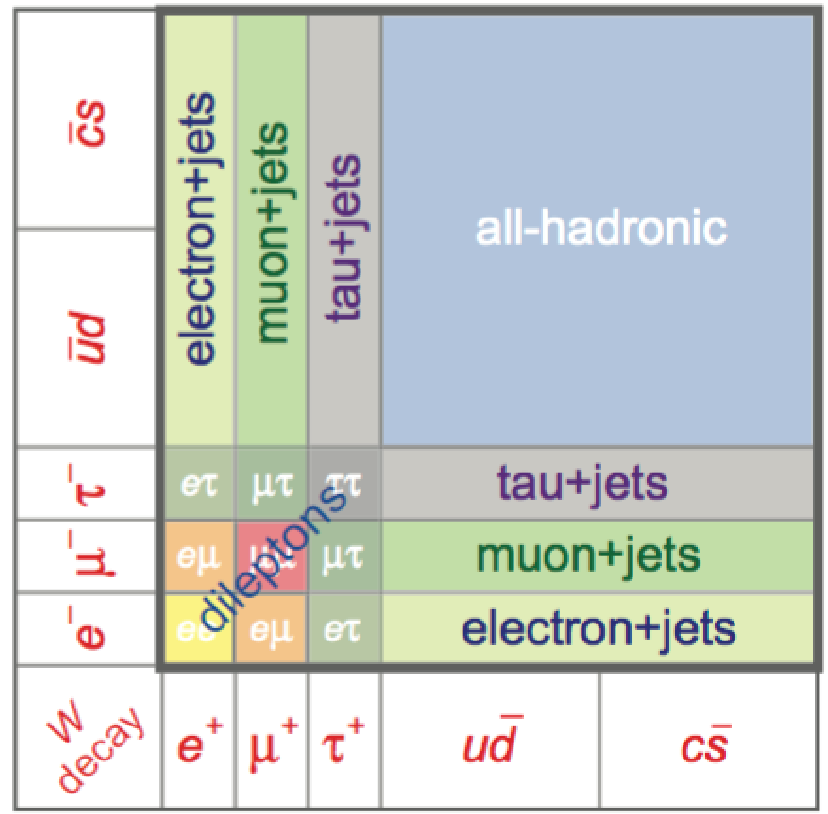


Single top (t-channel, Wt, s-channel)



- Dominant (99.8%) top decay t→Wb
  - Decay topologies dominated by W decay modes
  - W→eν, →μν, →τν, →qq (⇒jets)
- Nearly all object signatures are important
  - Electrons, muons (and taus)
  - Jets and b-tagged jets
  - Missing energy from neutrino(s)
    - … sorry, no diphoton decays (but tT+photon)

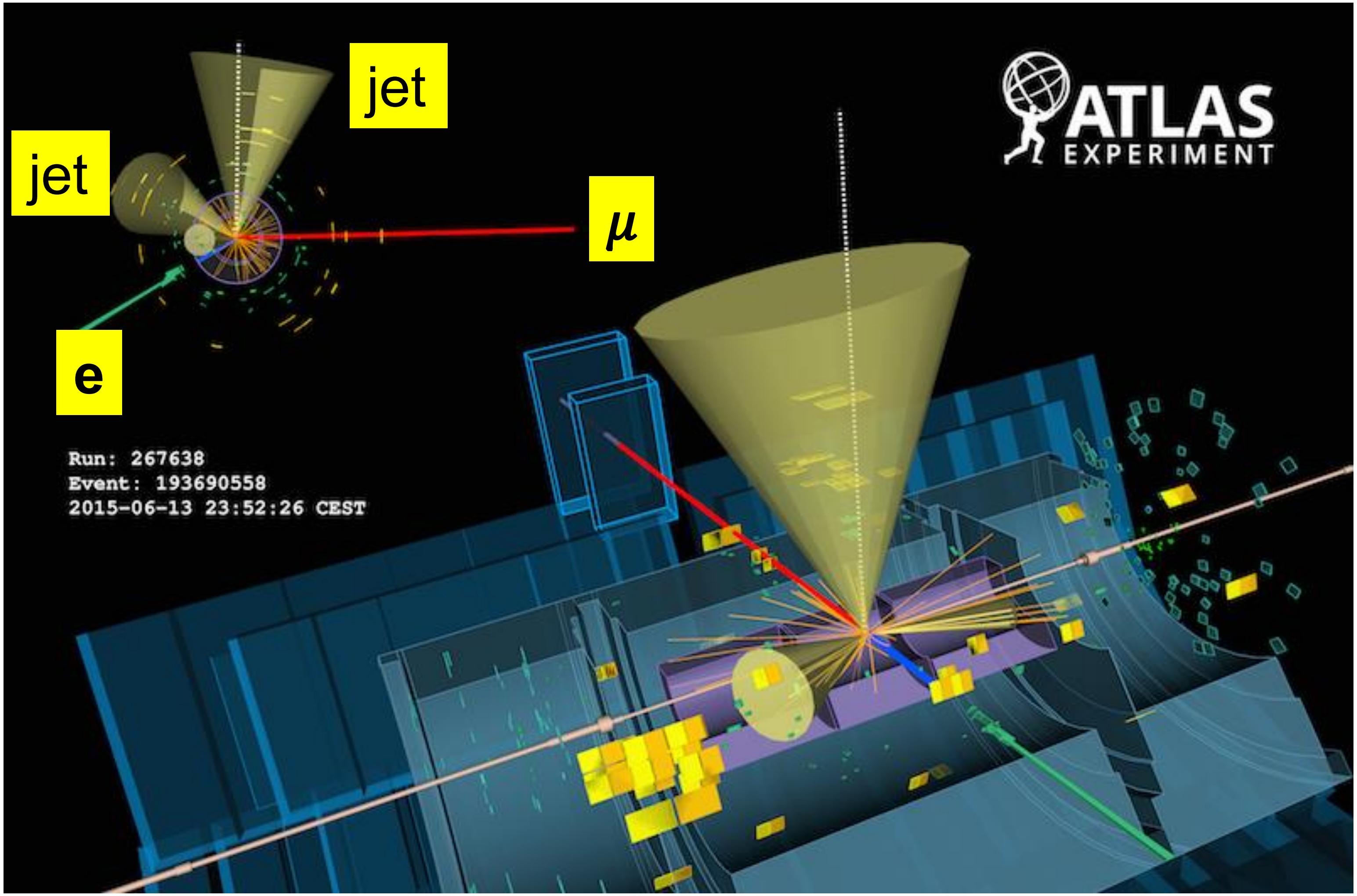
Final states for top pair (tT)



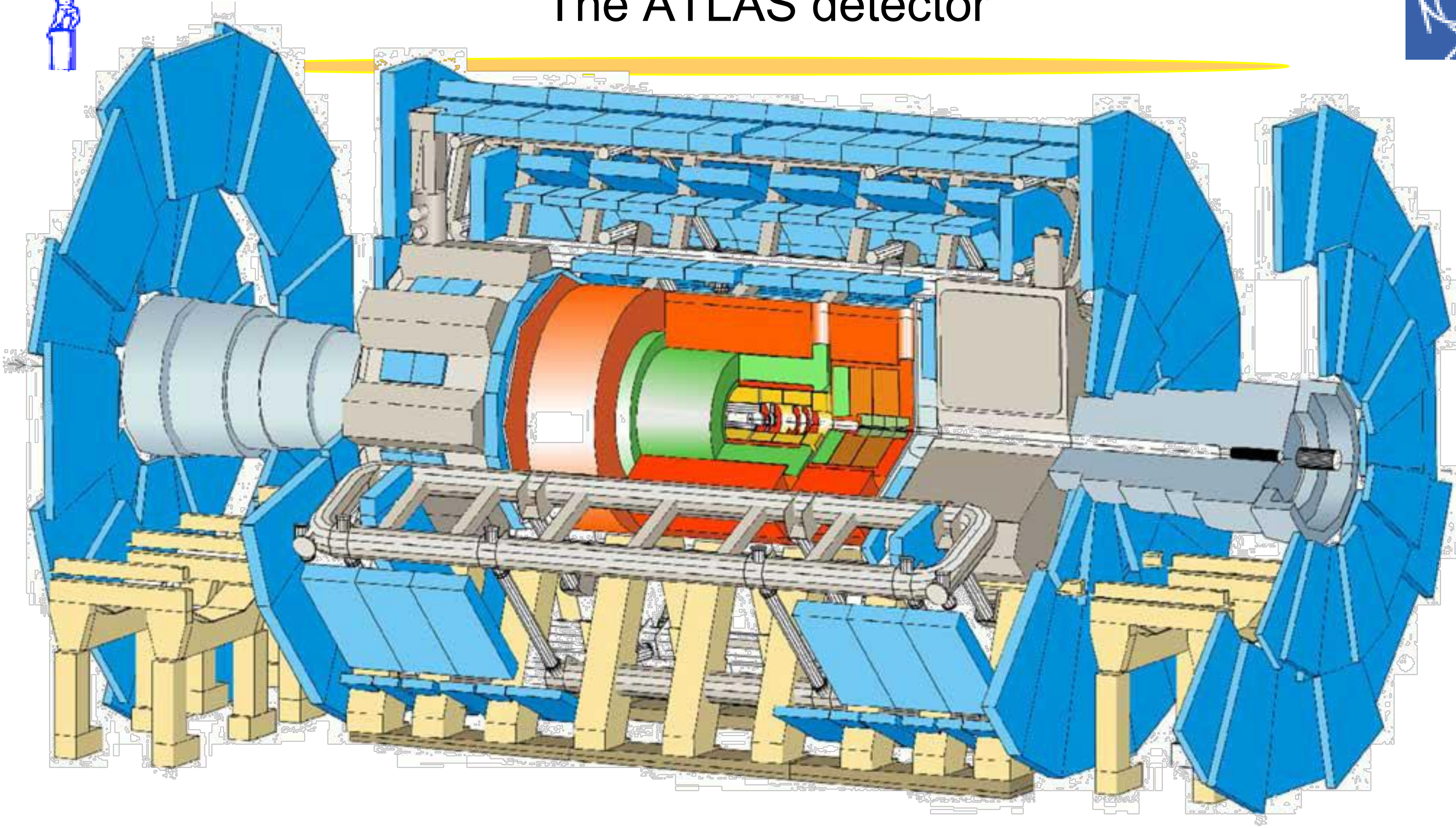
Objects for top physics …

Objects for top physics

- Or more realistically …early 13 TeV tT→e$\mu\nu\nu$bb with 2 b-tagged jets



The ATLAS detector

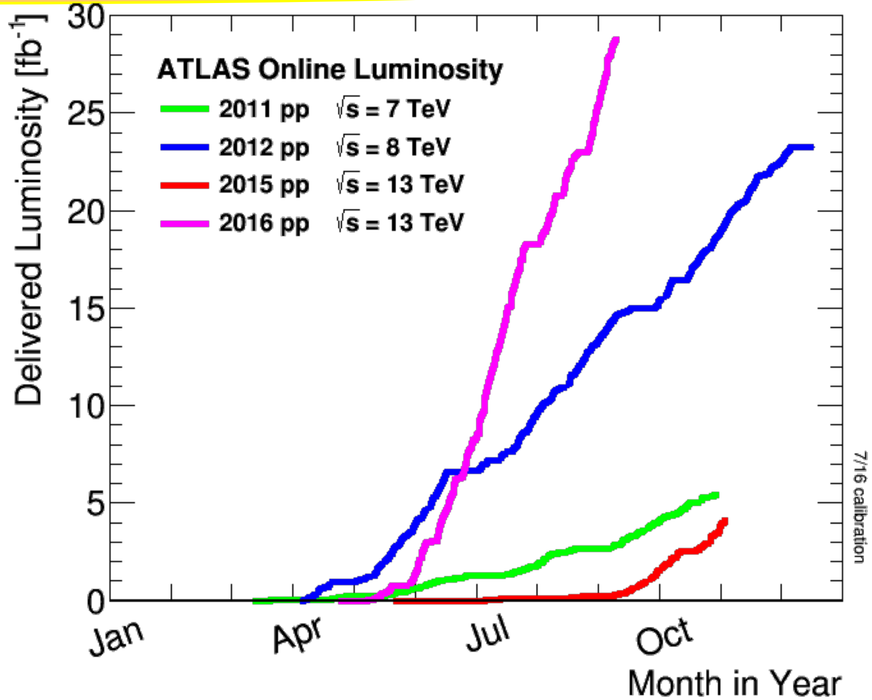


- Tracking detectors, EM and hadronic calorimeters and muon spectrometer
  - New Innermost B-layer (IBL) pixel layer at r=3.3 cm from beam for run-2

Data samples for top physics

- LHC has/is accumulating large top samples
  - Raw number of top pairs (tT) for each year:

Year	√s (TeV)	<μ>	L_{int} / fb^{-1}	σ(tT)/pb	N(tT)
2011	7	9	4.6	170	800k
2012	8	20	20.2	250	5M
2015	13	14	3.2	830	2.6M
2016	13	~25	> 25	830	> 21M



- Most results based on run-1 or 2015 data
  - Only a fraction of what we have now …
- 2016 sample could be 30-40 fb^{-1}
  - Excellent data quality so far, despite some challenges (e.g. toroid magnet)
  - 100+ fb^{-1} (approaching 100M top pairs) for run-2 total up to LS2
- How to best use this data for top physics?

ATLAS pp 25ns run: April-July 2016

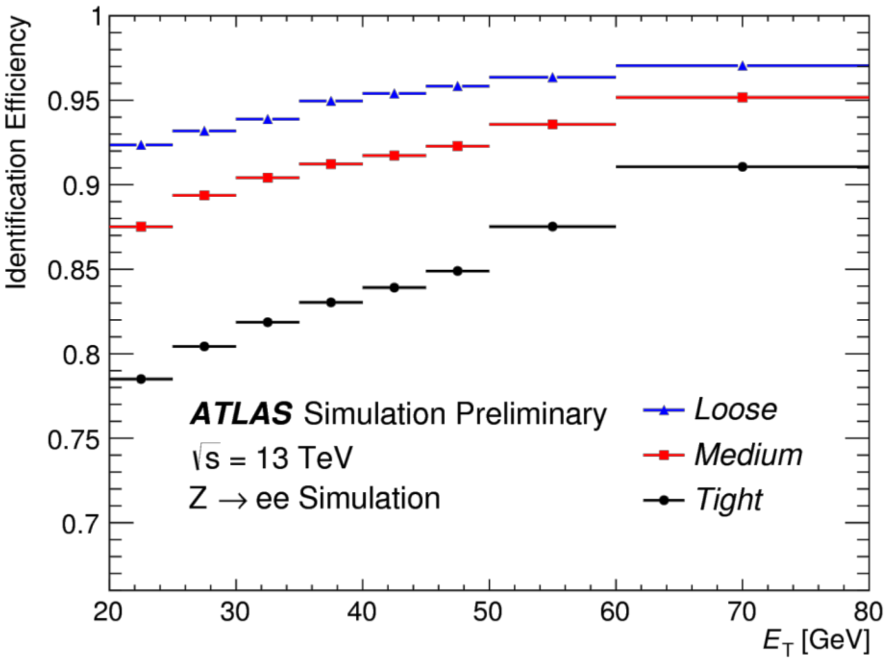
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
98.9	99.9	100	99.8	100	99.6	99.8	99.8	99.8	99.7	93.5

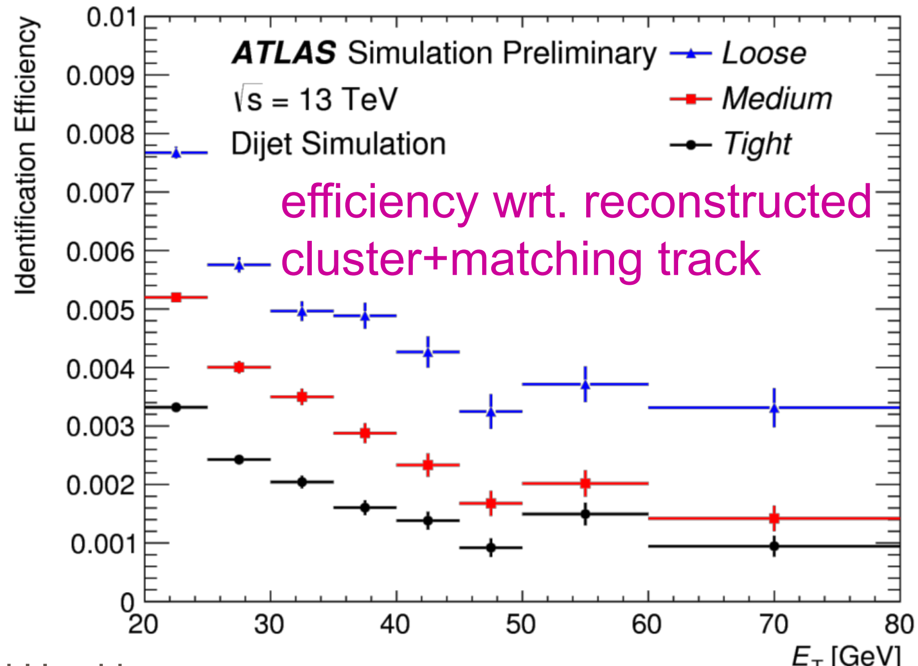
Good for physics: 91-98% (10.1-10.7 fb^{-1})

Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collisions with 25ns bunch spacing at √s=13 TeV between 28th April and 10th July 2016, corresponding to an integrated luminosity of 11.0 fb^{-1}. The toroid magnet was off for some runs, leading to a loss of 0.7 fb^{-1}. Analyses that don't require the toroid magnet can use that data.

Electron identification

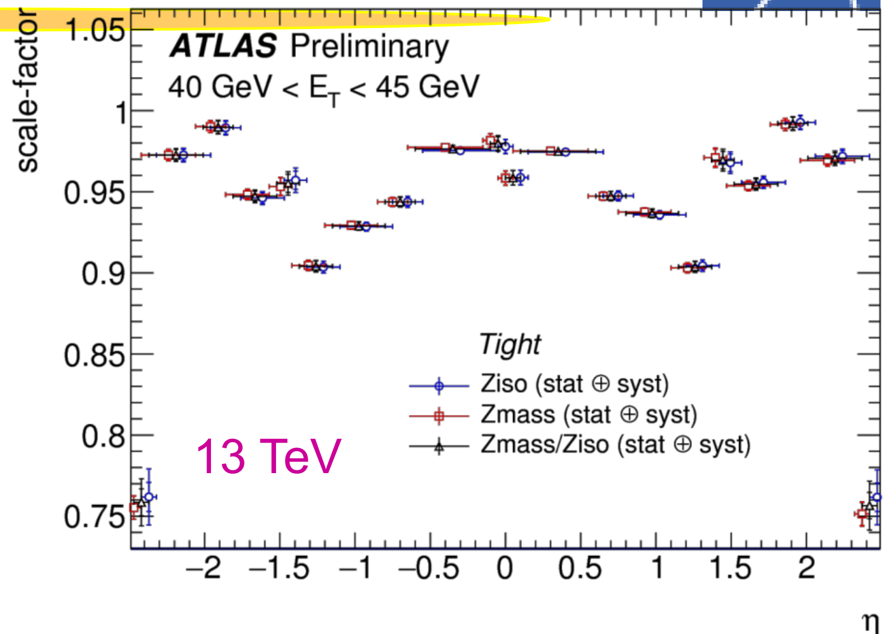
- Identified as EM calorimeter shower, spatially matched to ID track
  - Special track fit for electron candidates allowing for bremsstrahlung energy loss
  - Major backgrounds are misidentified hadrons and photon conversions
  - Require shower shape consistent with electron, E/p match, high-threshold hits in TRT detector, hit in first pixel layer (reject conversions)
- Use **medium** (dilepton) or **tight (**l+jets) likelihood based ID on cluster+trk
  - ID efficiencies of 80-95% (smaller at low p_T, high $|\eta|$), QCD jet rejection of $O(10^{-3})$

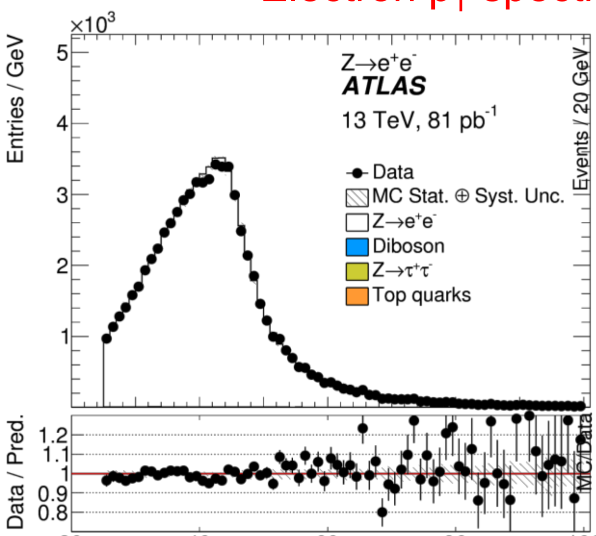


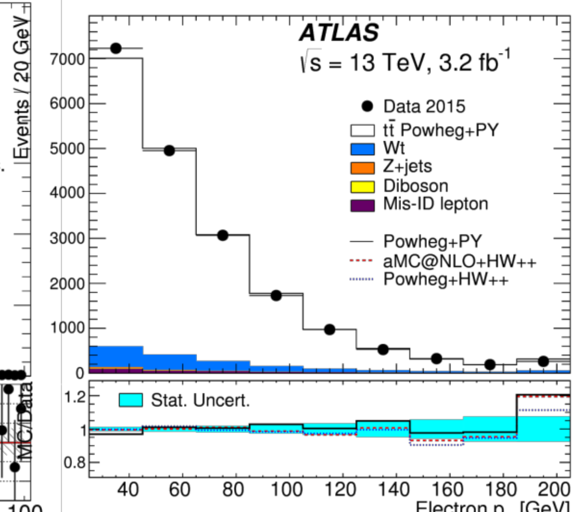


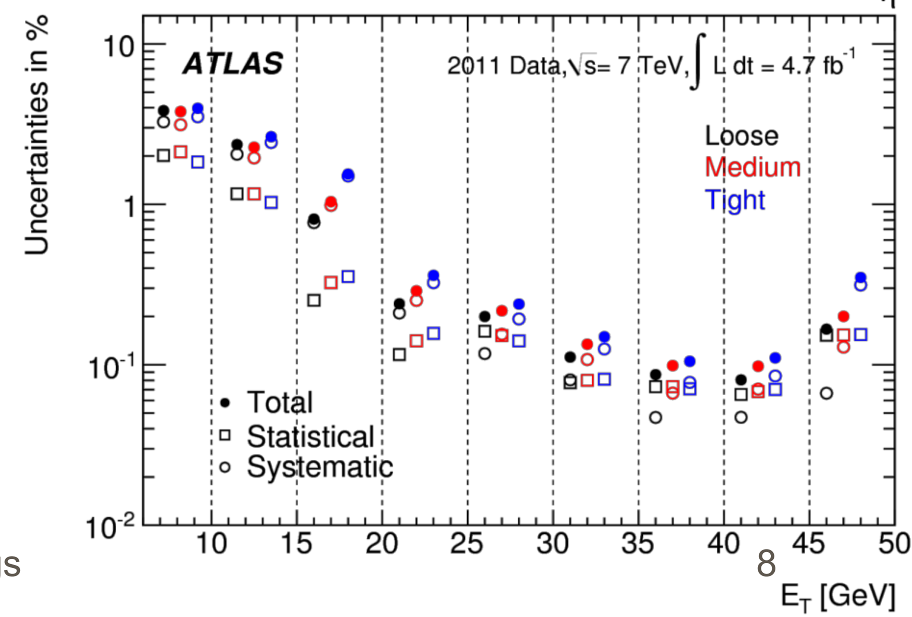
Electron efficiency measurements

- $Z \to ee$ (and $J/\psi \to ee$) used for **tag and probe**
  - One tightly-identified electron (tag), other just a track+calo cluster (probe), test ID req.
    - Z-mass requirement ensures probe sample is dominated by pure electrons
  - Efficiency ratio in data/MC –scale-factor
    - Typically within 5% of unity except in regions with high material
  - Uncertainties below 1% in relevant regions
    - Electron p_T spectra harder in $t \to e$ than $Z \to ee$



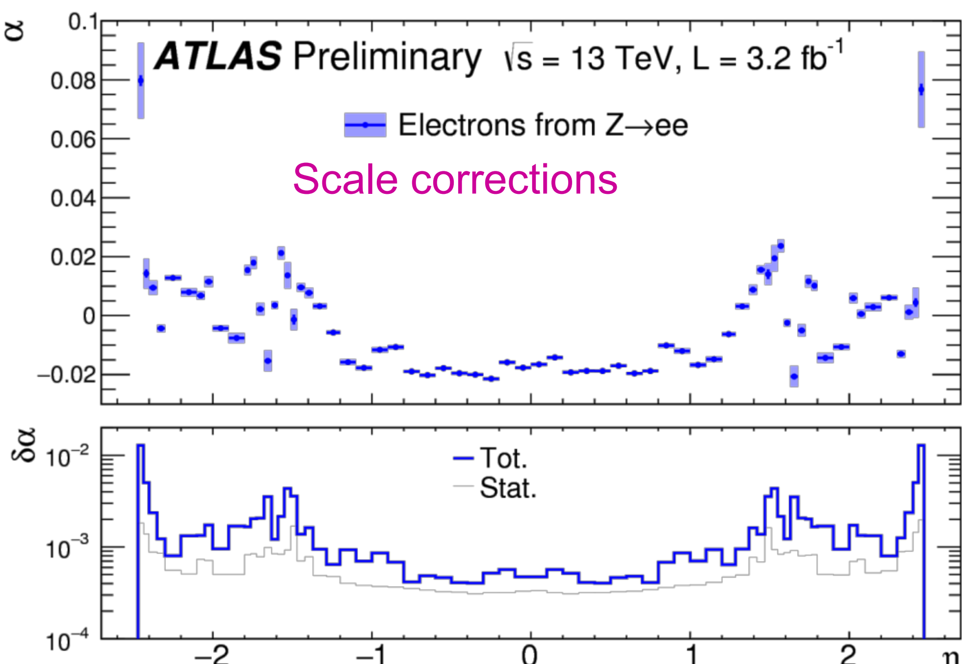


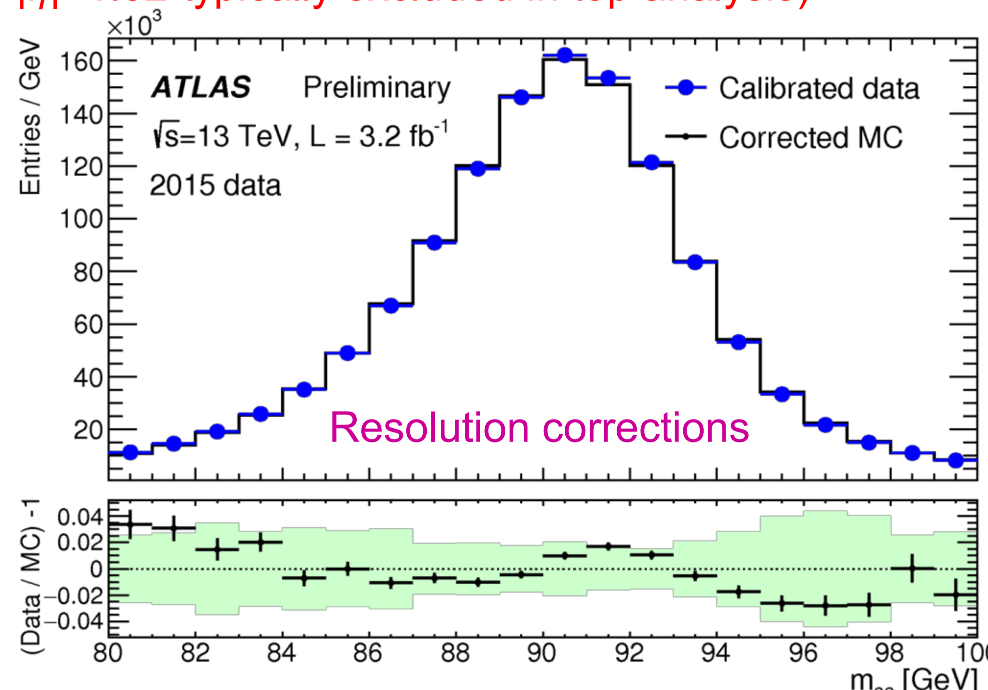




Electron energy calibration

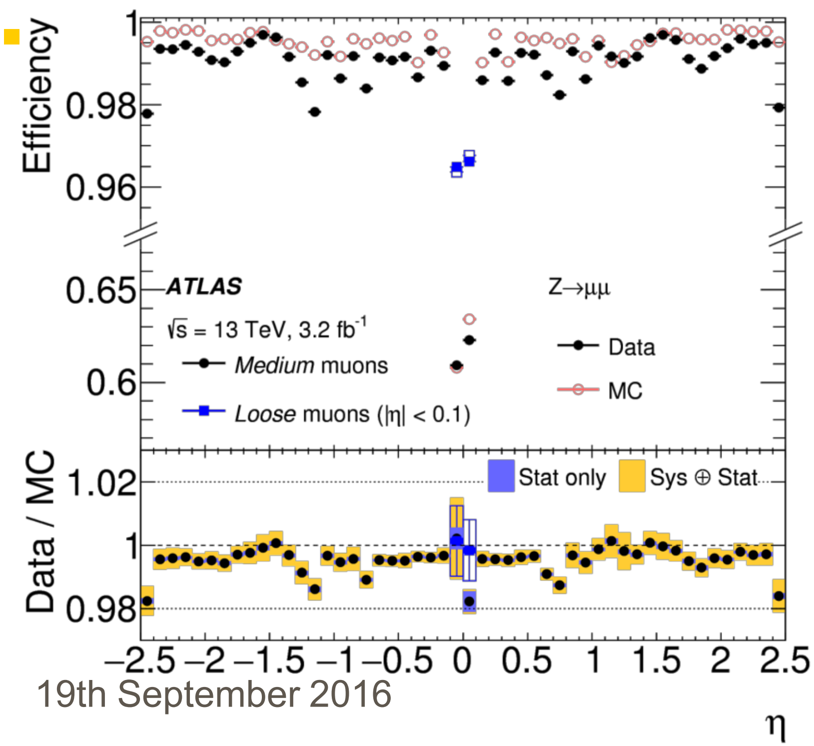
- Energy scale from Z→ee decays (known m_Z), validated with J/ψ→ee, Z→llγ
  - 'Bottom-up' cluster calibration based on simulation (+validated material model)
  - Final corrections using Z→ee in data: $E^{\text{data}} = E^{\text{MC}}(1+\alpha_i)$
    - Scale correction $\boldsymbol{\alpha}$ derived from MC template fits to data in bins of η , together with corrections to resolution constant term
  - Final scale corrections up to a few %, energy scale uncertainties < 10^{-3}
    - Except in barrel-endcap transition (1.37<|η|<1.52 typically excluded in top analysis)

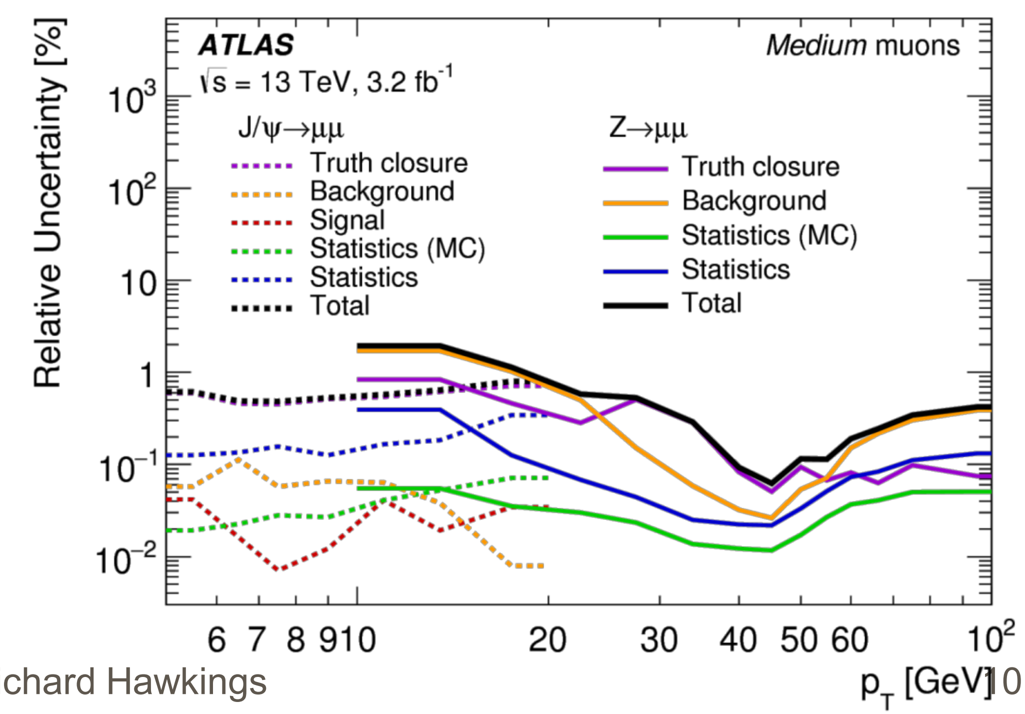




Muon identification and efficiency measurement

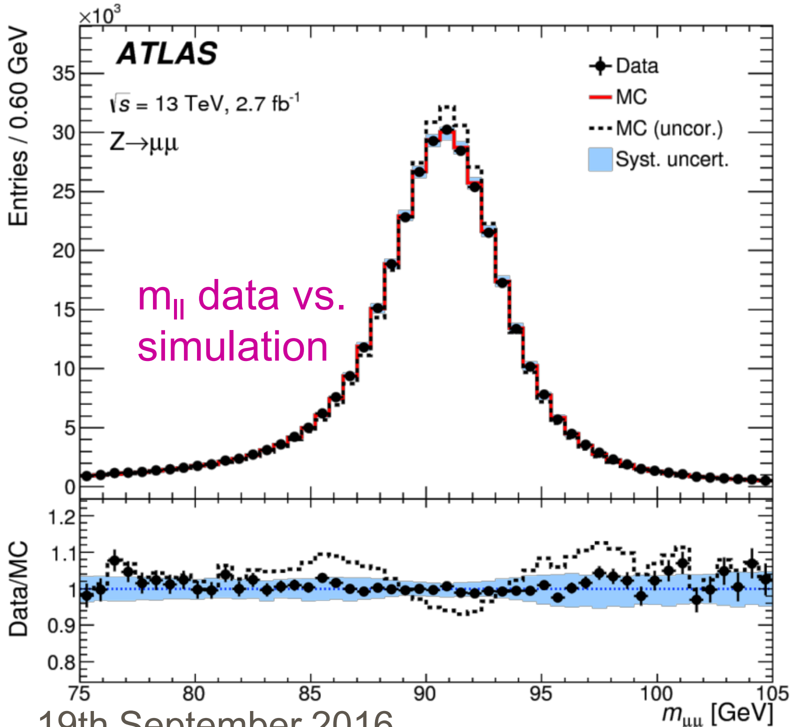
- Muon candidates from independent tracks found in inner detector (ID) and muon spectrometer (MS), combined with global track fit
  - ‘Medium’ requirements typically used – compatibility of q/p of ID and MS tracks, together with hit/quality requirements on individual tracks
  - Main backgrounds from π/K decays in flight, hadronic ‘punch through’ calorimeter
- Efficiencies measured using T&P with $Z\rightarrow\mu\mu$ and $J/\psi\rightarrow\mu\mu$, as for electrons
  - Typically above 98% for medium muons; uncertainties 0.1-1% in relevant p_T range

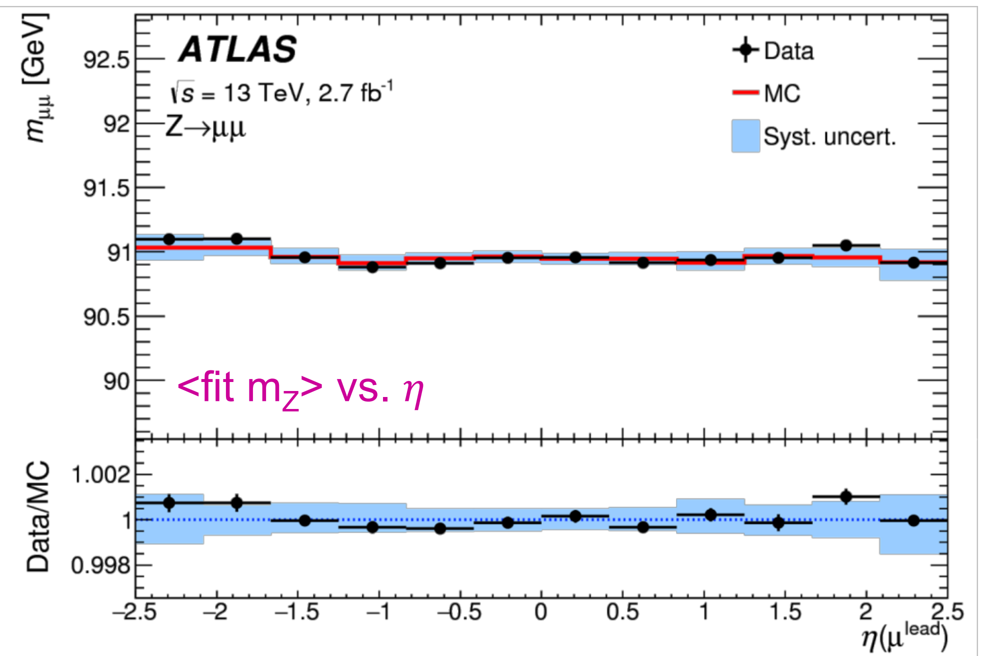




Muon momentum calibration

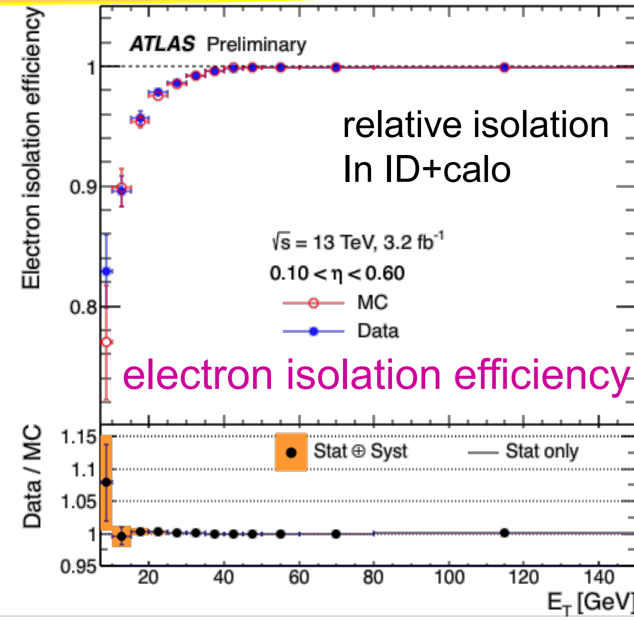
- Muon momentum scale / resolution depend on ID alignment, chamber drift time calibration and alignment, magnetic field, knowledge of material…
  - Final absolute calibration from $Z\to\mu\mu$ and $J/\psi\to\mu\mu$ mass distributions in data/MC
    - Template fits in different (η,ϕ) regions of the detector
    - Adjust scale and resolution parameters for ID and MS contributions separately
  - Final scale uncertainties at or below 10^{-3} over full rapidity range

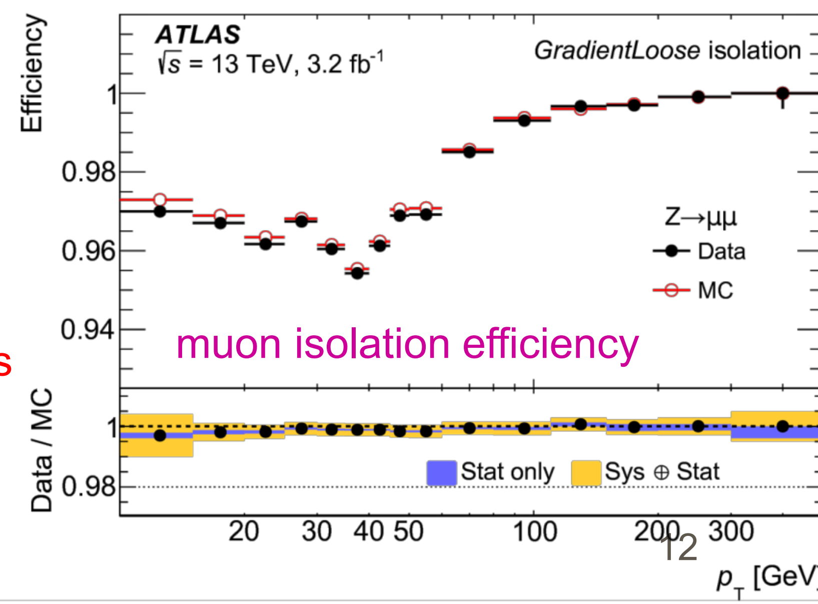




Lepton isolation

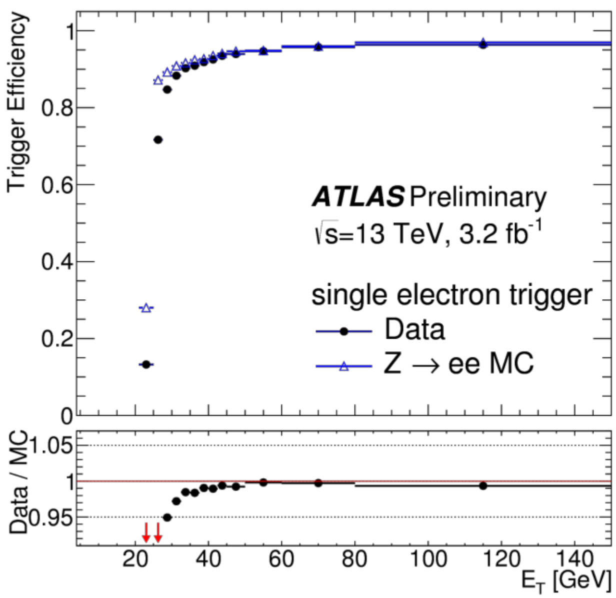
- Leptons isolated from nearby hadronic activity
  - Reduces background e.g. from b→l, c→l
    - Calorimeter: energy in $\Delta R<0.2$ cone around lepton
    - Tracking: sum of track p_T in a variable-sized cone dependent on lepton p_T
  - Background is more significant for low p_T leptons
    - Cut on relative isolation p_T^{cone}/p_T^{lepton} or tune cuts as function of p_T
    - Typical efficiency 95→99% for 25-60 GeV p_T
  - Efficiency measured on data with Z→ll T&P
  - Typically also require leptons to be separated by $\Delta R>0.4$ from all reconstructed jets
- Top environment 'busier' than Z→ll
    - Parameterise efficiency dependence on nearby jets (with Z+jet events), or …
    - …Measure 'in-situ' in data tT events by relaxing cuts
  - Can reduce MC modelling uncertainty – MCs predict different hadronic activity near lepton

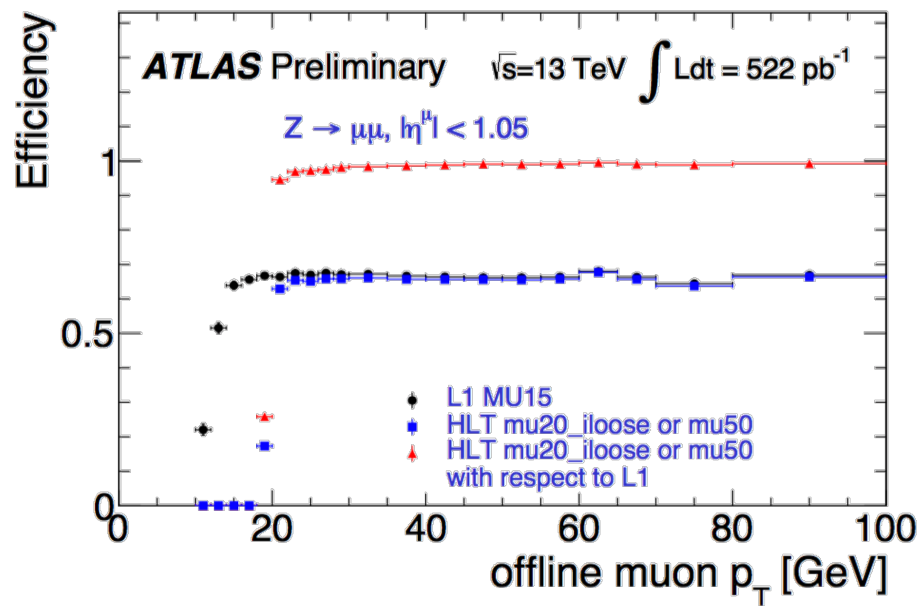




Lepton triggers

- ATLAS top analyses (even dileptons) typically use **single lepton** triggers
  - Thresholds fully efficient for electrons and muons with p_T>25 GeV
- Efficiencies from data using Z→ll T&P for leptons passing offline selection
  - Electrons 90-95% (turn-on at low p_T), muons ~70% (barrel), 85% (endcaps)
  - Dilepton tT can be triggered by either lepton: ~99% per-event efi, low systematics





  - Cannot maintain 25 GeV thresholds for LHC luminosity L>1 10^{34} $cm^{-2}s^{-1}$
    - Raise p_T threshold by few GeV, or use dilepton or lepton+jet/E_T^{miss} trigger for full 2016 data
    - See **poster** by Michele Faucci Gianelli for more details …

Leptons in action – related uncertainties

- Relative uncertainties (%) on $t\bar{T}\rightarrow e\mu$+ b-jets incl. cross-section measurement

Source / (%)	7 TeV	8 TeV	13 TeV
Electron efficiency	0.13	0.41	0.3
Electron scale/res	0.22	0.51	0.2
Electron isolation	0.6	0.3	0.4
Muon efficiency	0.30	0.42	0.4
Muon scale/res	0.14	0.02	<0.05
Muon isolation	0.4	0.2	0.3

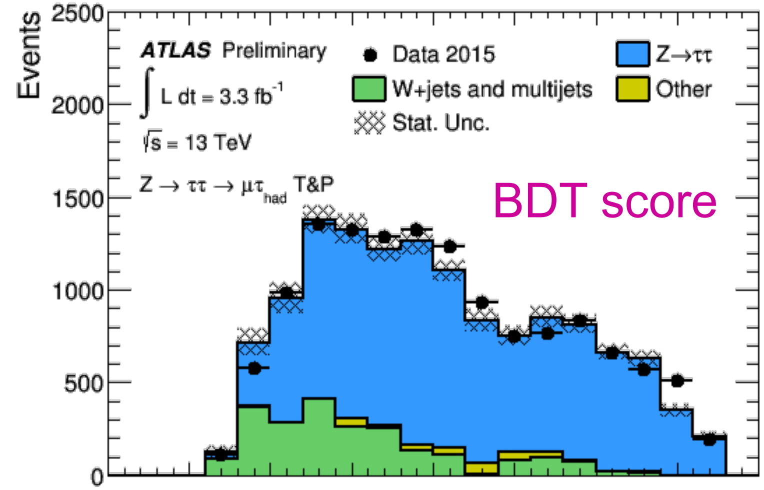
- All below 1%, run-1 and run-2 comparable
  - Electrons worse at 8 TeV due to use of non-final material model
- Uncertainties on top quark mass due to leptons (efficiencies + scale/res^{ln})

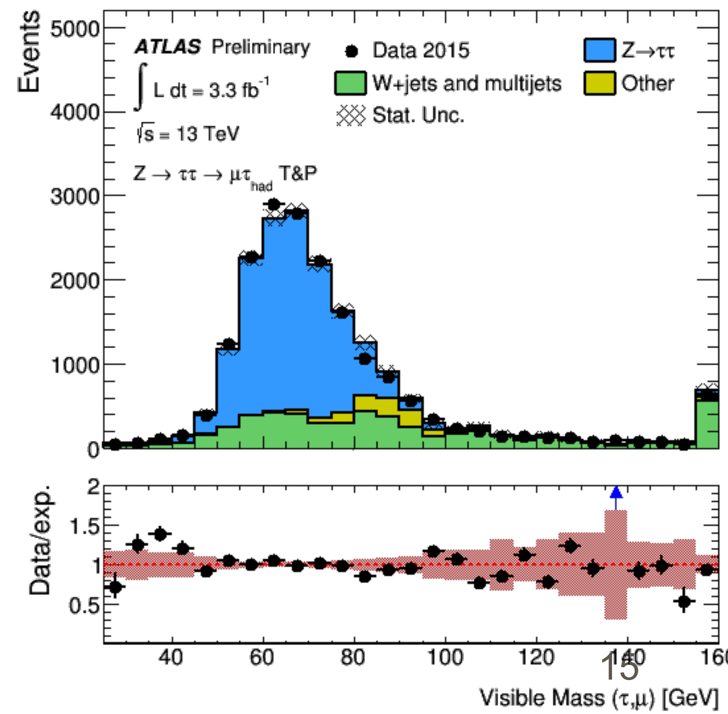
	l+jets (7 TeV)	ll (7 TeV)	ll (8 TeV)
Lepton uncertainty (GeV)	0.04	0.13	0.14

  - Sub-leading uncertainty source in top mass (lepton scale uncertainties <10^{-3})

Beyond e and μ - tau leptons

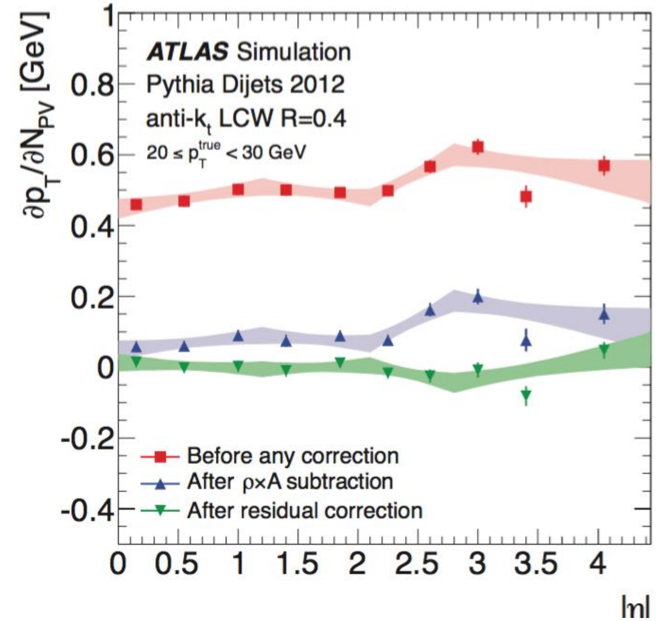
- τ-identification much harder than for e and μ
  - Leptonic τ decays (BR($\tau \rightarrow e/\mu+\nu\nu$)=35%) give extra contributions to $t \rightarrow W \rightarrow e/\mu$ channels
    - Mainly at low lepton p_T, treated as e/μ signal
- Hadronic τ gives narrow jet with low associated track multiplicity (1, 3 prong)
  - Use BDT to separate from hadronic jets
    - Shower shapes, track isolation cones, track momentum fractions, impact parameters, …
- Efficiency and energy scale based on $Z \rightarrow \tau\tau$ with one $\tau \rightarrow \mu\nu\nu$ and one hadronic τ
  - Uncertainties of 2-4% achieved with run-1 data
  - Backgrounds are topology-dependent – have to evaluate 'in-situ' with control regions (OS/SS)
- Results generally not competitive with e/μ
  - Important for new physics searches
    - e.g. charged Higgs: $t \rightarrow H^+ \rightarrow \tau$ vs. $t \rightarrow W \rightarrow e/\mu/\tau$



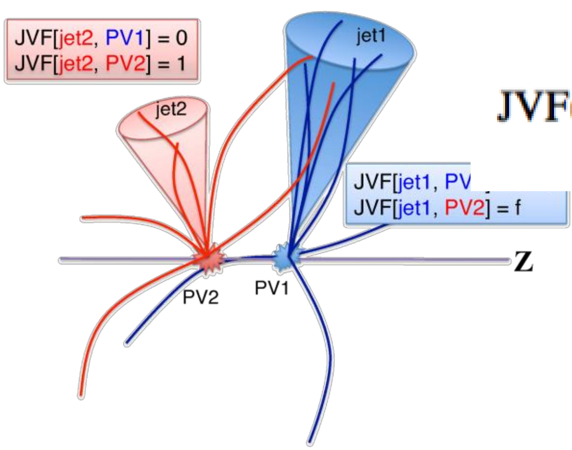


Jet reconstruction and pileup suppression

- Crudely, outgoing quarks and gluons reconstructed as jets of particles in detector
  - ATLAS uses R=0.4 anti-k_T jets formed from topological calorimeter clusters
  - Calibrated from MC, with data-based correction
- Pileup adds energy to each measured jet
  - Subtract using 'jet-area' corr^n: $p_T^{corr} = p_T^{jet} - \rho \times A^{jet}$
    - p_T density ρ from median of k_T jets in $|\eta|<2$
  - After residual corr^n of N_{pv},$<\mu>$ effects, flat dp_T/dN_{pv}

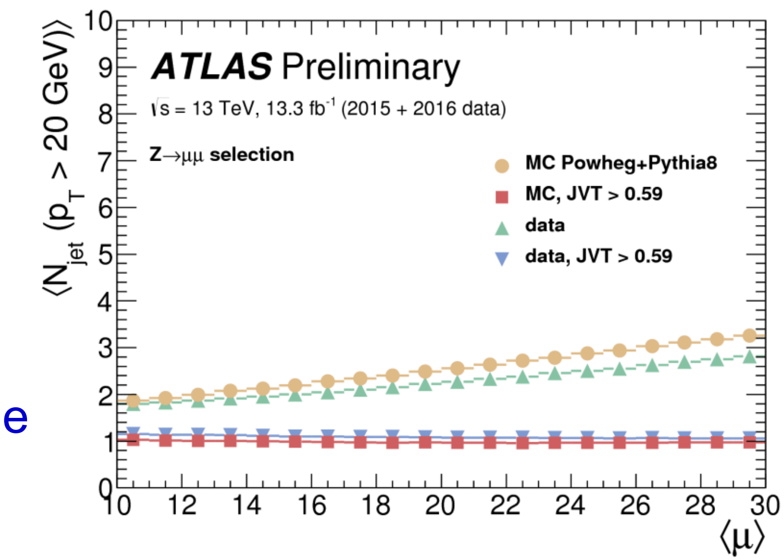


- Remove jets from pileup with 'jet vertex fraction'



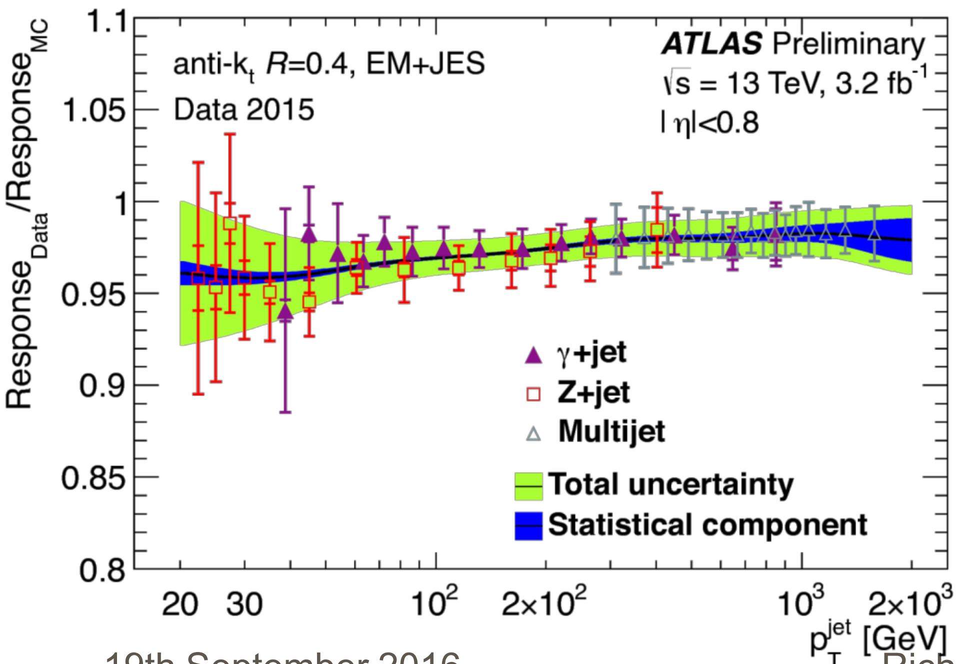
$$\mathrm{JVF}(\mathrm{jet}_i, \mathrm{PV}_j) = \frac{\sum_m p_T(\mathrm{track}_m^{\mathrm{jet}_i}, \mathrm{PV}_j)}{\sum_n \sum_l p_T(\mathrm{track}_l^{\mathrm{jet}_i}, \mathrm{PV}_n)}$$

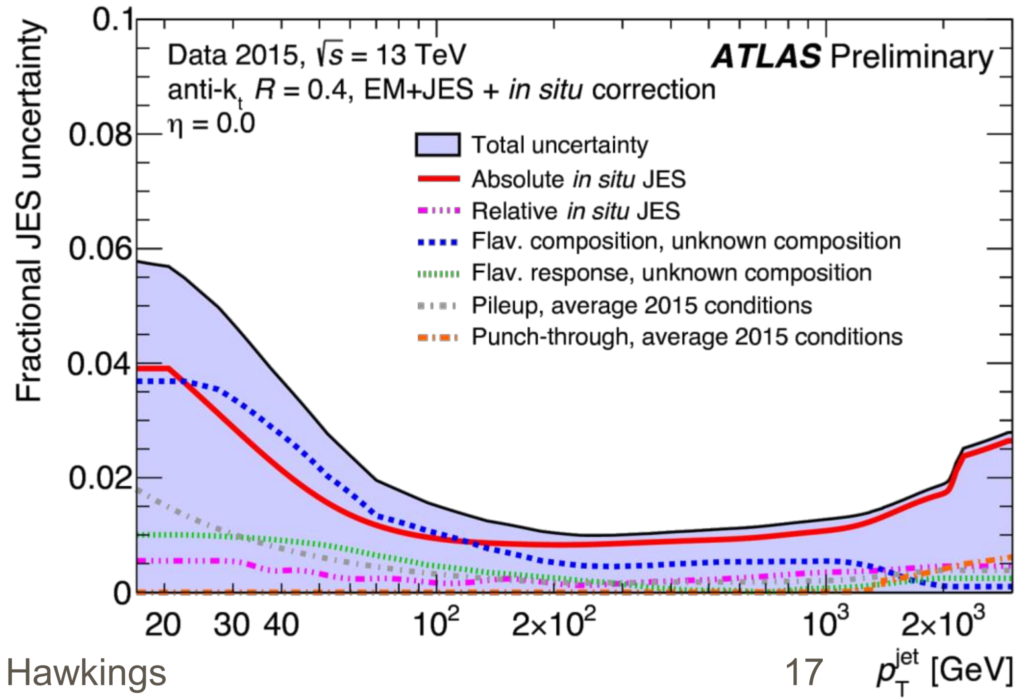
- Enhanced 'jet vertex tagger' used in run-2
- Jet multiplicity in $Z\rightarrow\mu\mu$ stable vs pileup $<\mu>$



Jet energy scale calibration

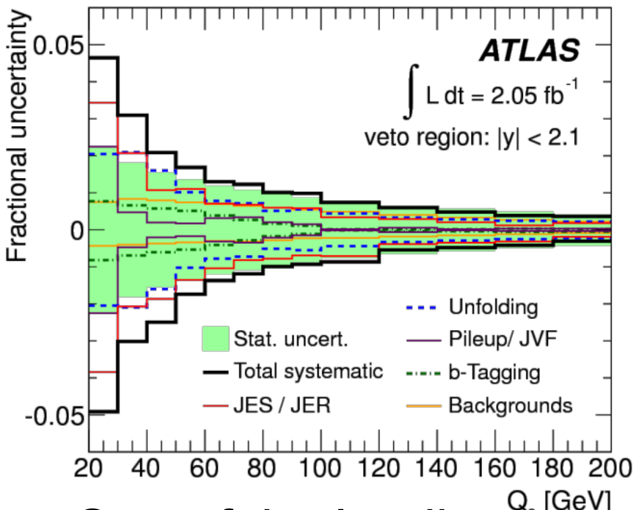
- Jet energy scale calibration adjusted with in-situ corrections from data
  - Use p_T balance in photon+jet and Z(→ee)+jet events to calibrate against well-known EM scale (from Z mass)
  - Multijet events (1 high p_T recoils against 2 or more lower p_T) to extend to higher p_T
- Energy scale known to e.g. <2% at p_T≈100 GeV in 2015, worse for low p_T
  - Almost factor 2 better in final run-1 calibration
  - Also significant dependence on jet flavour composition(quark, gluon, b-jet)



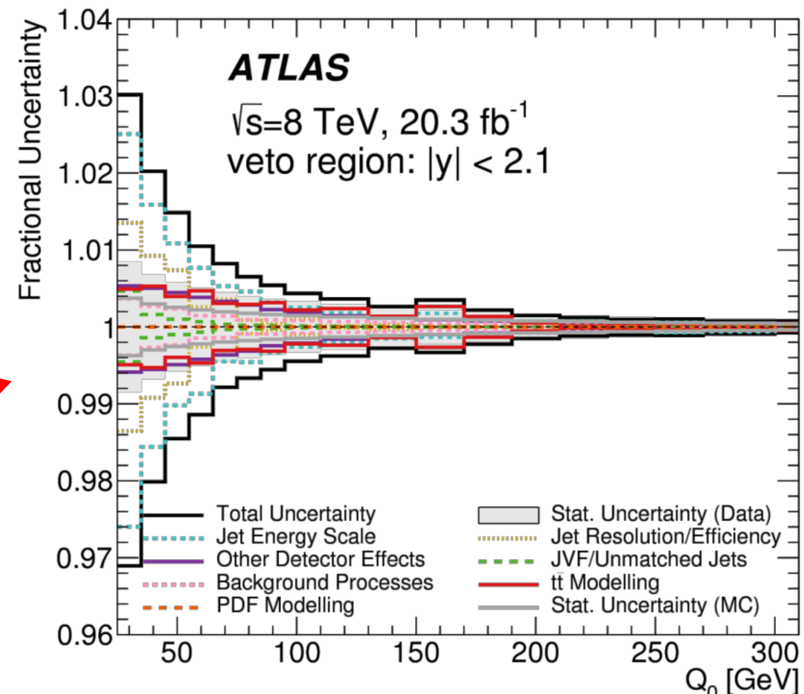


Impact of jet uncertainties

- Jet energy scale uncertainty expressed as O(20) uncorrelated components
  - With different dependencies on jet p_T and $|\eta|$
- Often leading detector-related uncertainty
  - E.g. in gap fraction measurement in $e\mu bb$



Improvement at 8TeV from jet area corr^n, despite higher pileup

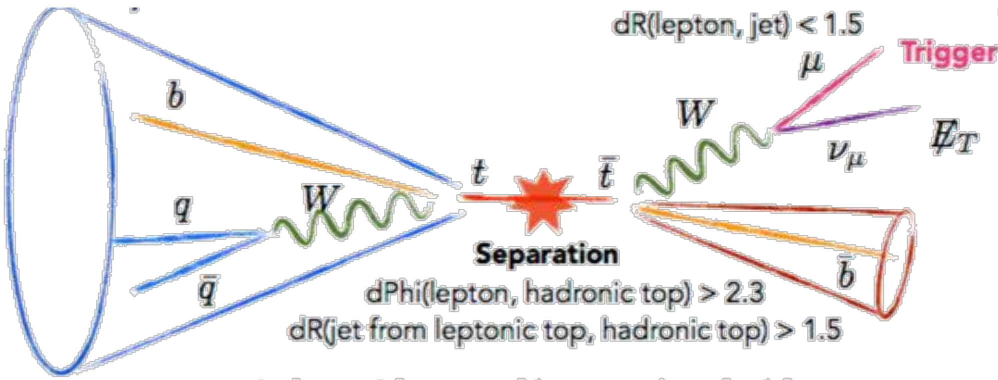


- One of the leading uncertainties in top mass measurements
  - Including effect of in-situ W mass constraint for l+jets (overall energy scale factor)

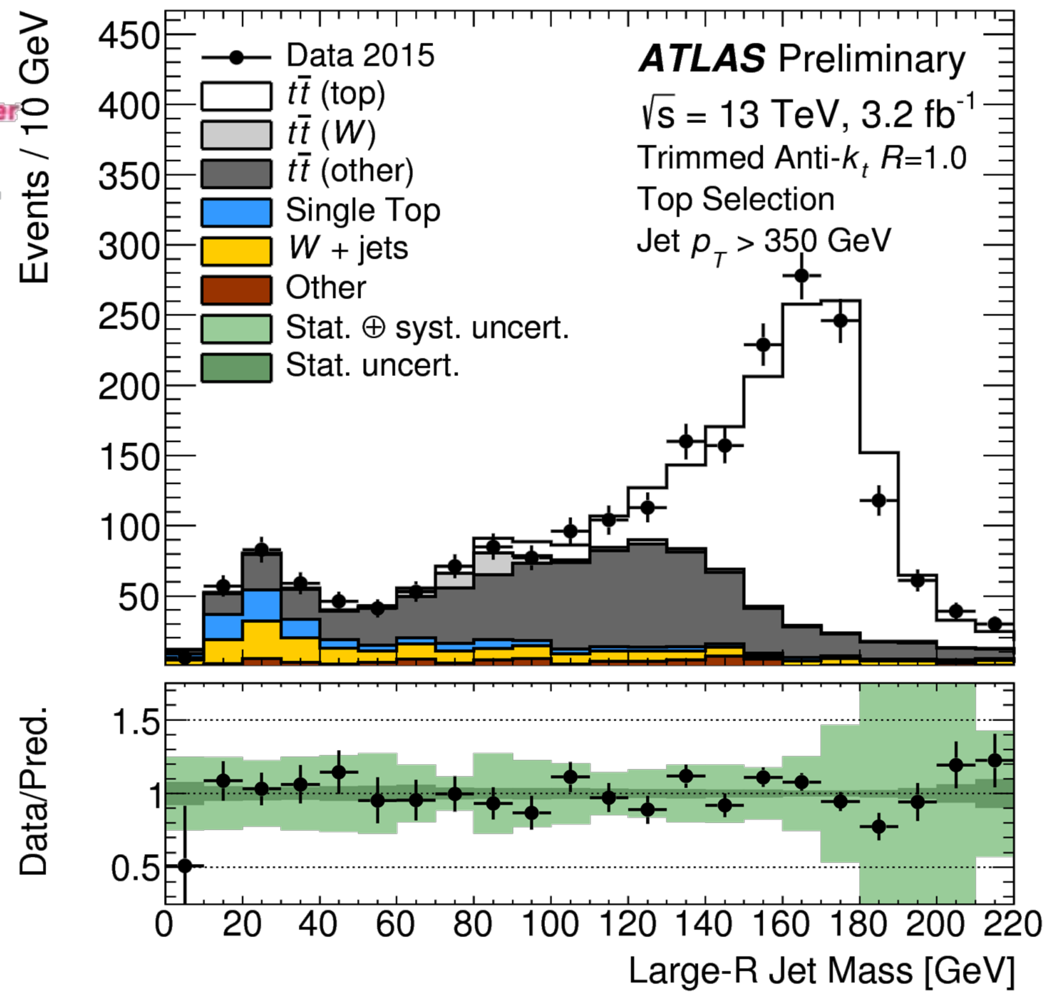
Uncertainty (GeV)	l+jets (7 TeV)	ll (7 TeV)	ll (8 TeV)	had (8TeV)
Jet energy scale	0.58	0.75	0.54	0.60
Relative b to light energy scale	0.06	0.68	0.30	0.34

Large-R jets for boosted topologies

- At high m(tT) / p_T(t) top decays are boosted: 3 jets from t→bqq merge

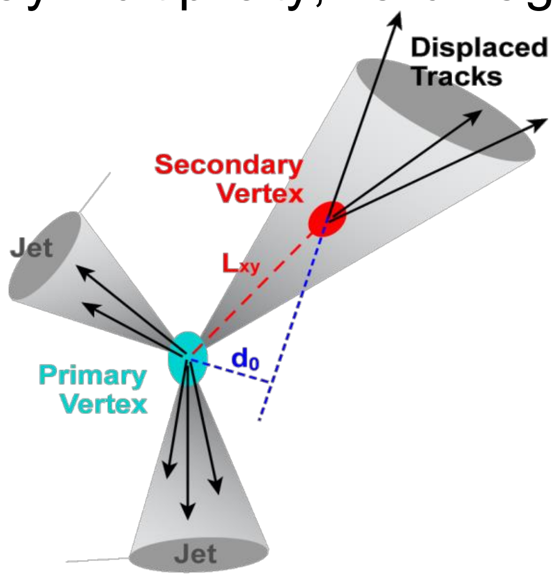


  - Use large-R (R=1.0) jets to capture all the top decay products into one jet
  - Use jet 'trimming' to remove soft contributions from pileup
- Selection with R=1.0 p_T>350 GeV jet + lepton, E_T^{miss} and R=0.4 b-tagged jet
  - Clear peak in large-R jet mass at m_{top}
  - Refine with jet substructure variables
- Boosted tT x-sec, BSM searches, …

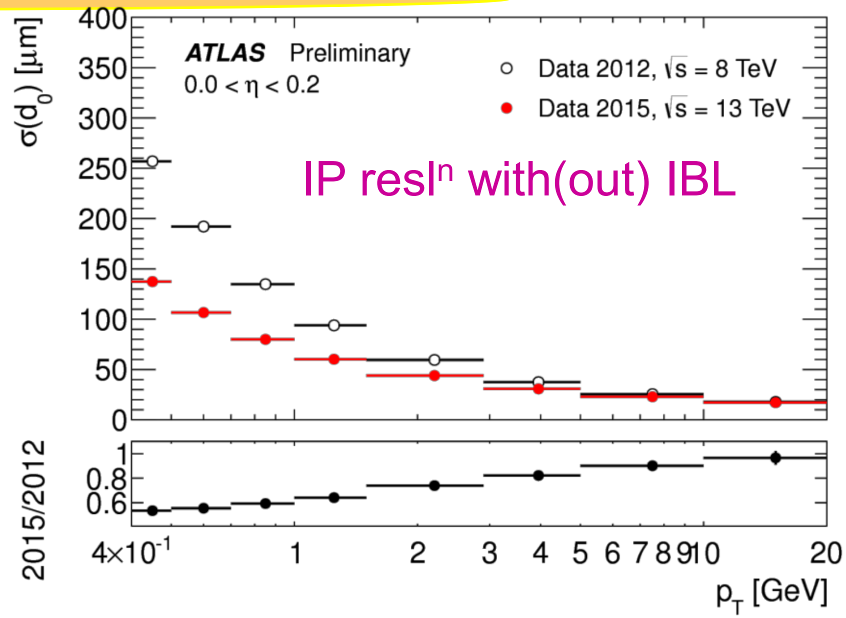


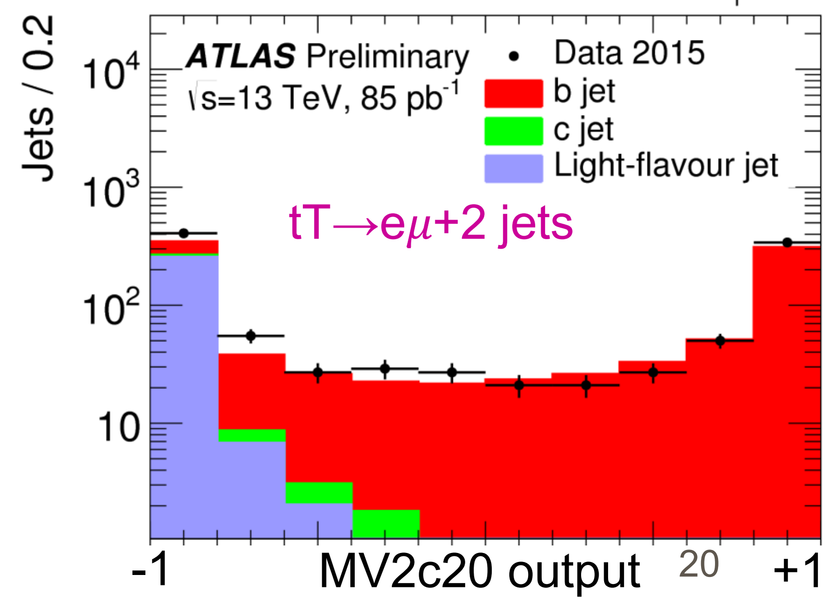
Tagging jets with b-flavour

- With BR(t→Wb)=99.8%, b-tagging is an important tool for top physics
  - Select tT and single top events
  - Separate b-jets from W→qq and radiation
- Relies on b lifetime (~mm decay length), high mass, decay multiplicity, hard fragmentation



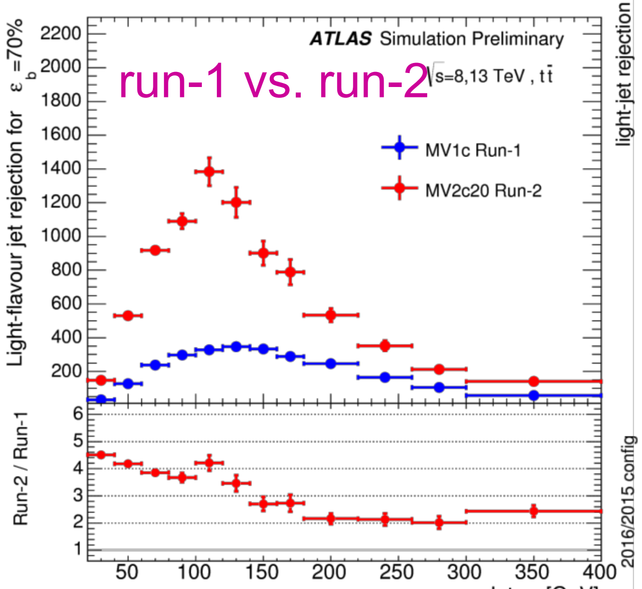
- Good impact parameter resolution is key
  - Information from various algorithms combined in an MVA (neural network, now BDT @ run2)

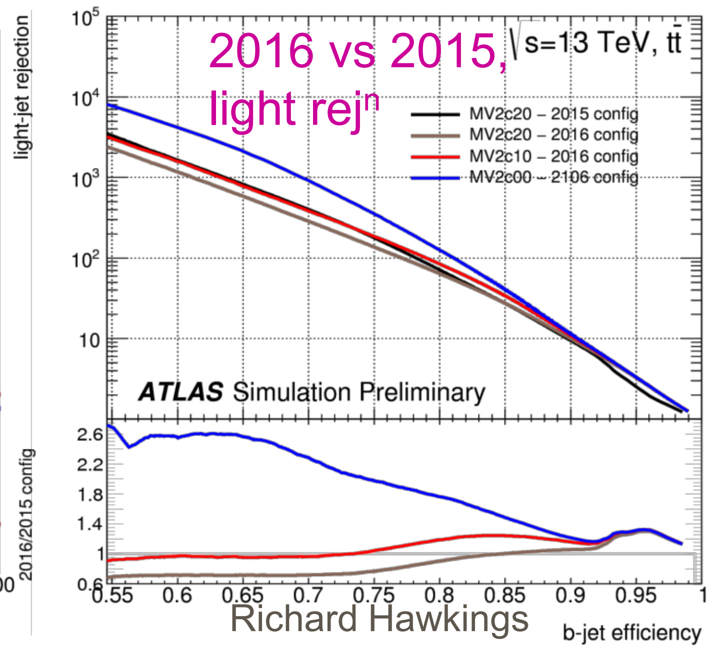


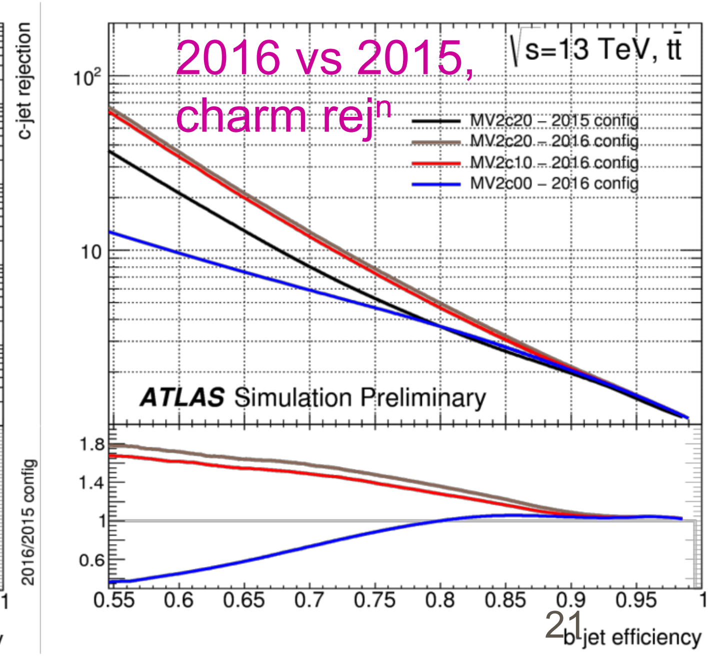


b-tagging performance improvements for run-2

- Comparison of run-1 and run-2 (incl. IBL) detectors and software, 13 TeV tT
  - Light jet rejection with constant 70% b-tagging efficiency for all jet p_T
  - Gains at low p_T from IBL (better IP resl^n) and high p_T from tracking/b-tag algorithms
    - Most physics analyses benefitted by moving from 70% to 77% b-tag efi. working point
- Further improvements to tracking, b-tagging and BDT training for 2016
  - Trade light quark for c rejection by modifying background mixture in training
    - MV2cxx: training with xx% of charm jets in background sample, 100-xx% light jets
    - MV2c10 (**red points**) default in 2016, cf. MV2c20 in 2015 (black): ~40% better c-rej

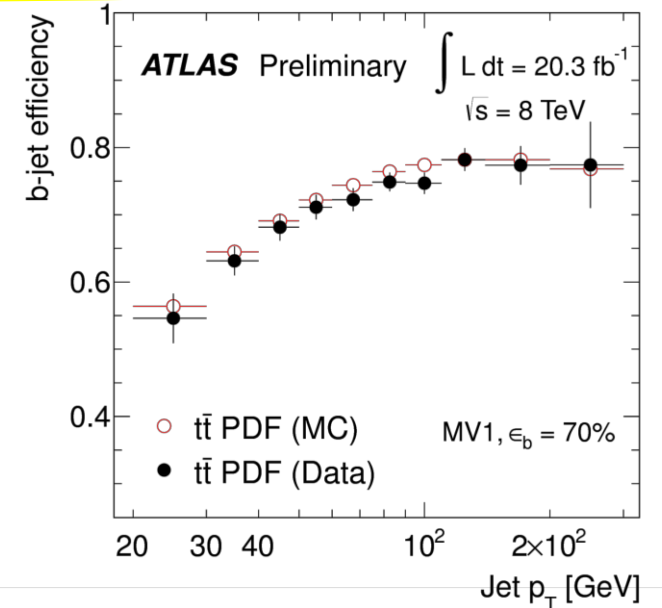


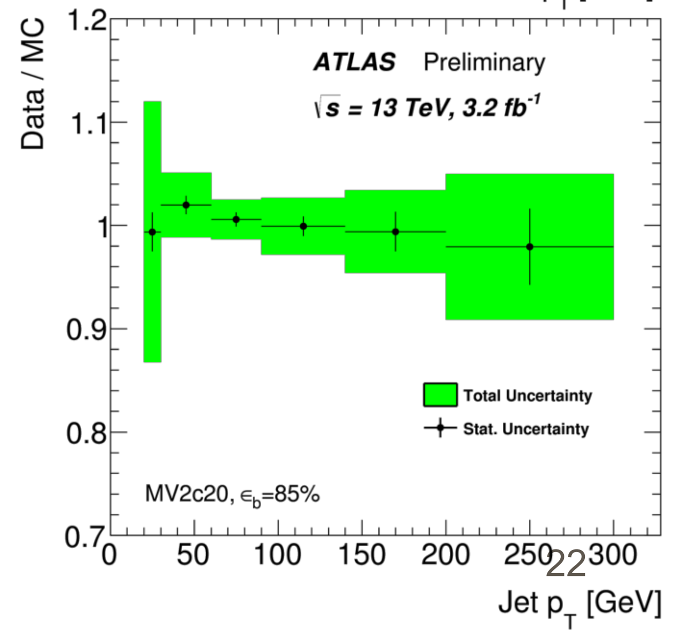




b-tagging in top events

- Precise b-tag efi. calib. from tT→ll$\nu\nu$bb events
  - Clean well-understood topology, rich in b-jets
  - Tag and probe (giving 1 unbiased b-jet)
  - 'PDF' likelihood calibration, exploiting all jets in ll+2 and ll+3 jet events
  - Precision of 2-3%in 50-100 GeV p_T range, limited by JES and tT modelling
    - Data/MC differences expressed as scale factors (≈1)
    - Uncertainties and correlations expressed with 10 eigenvector components, similar to JES
  - Charm and light jet calibration from D*, W+c and light jet events (20-50% uncertainties)
- Important uncertainties for high p_T top analyses, tT+heavy flavour, top mass measurement
  - E.g. 0.50 GeV for 7 TeV l+jets top mass
  - Mismodelling in b-tagging efficiency vs. p_T can bias b-tagged jet p_T distribution and hence m_{top}



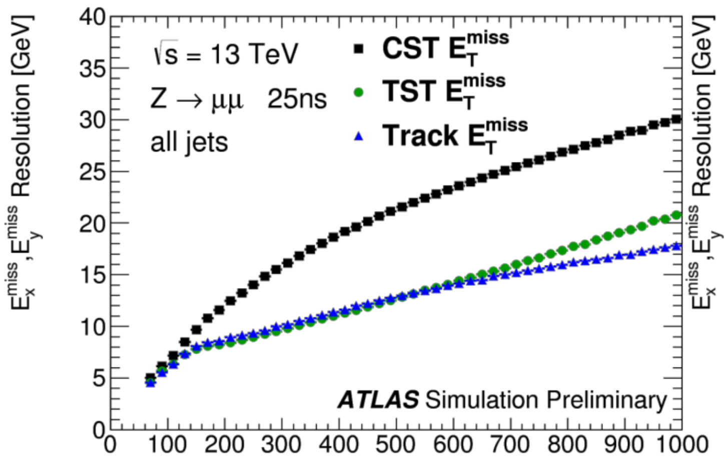


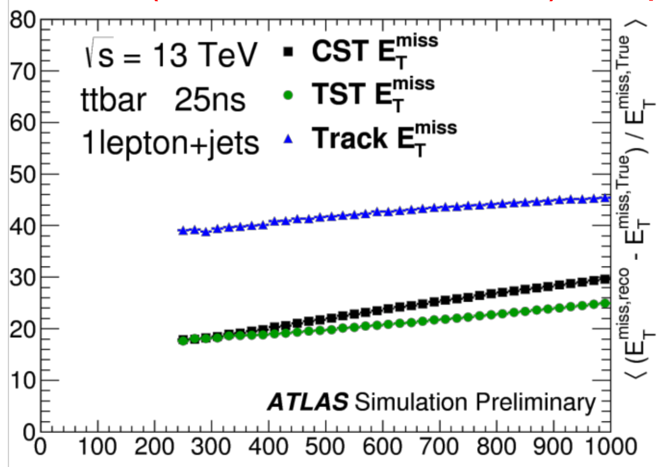
Missing transverse energy – capturing neutrinos

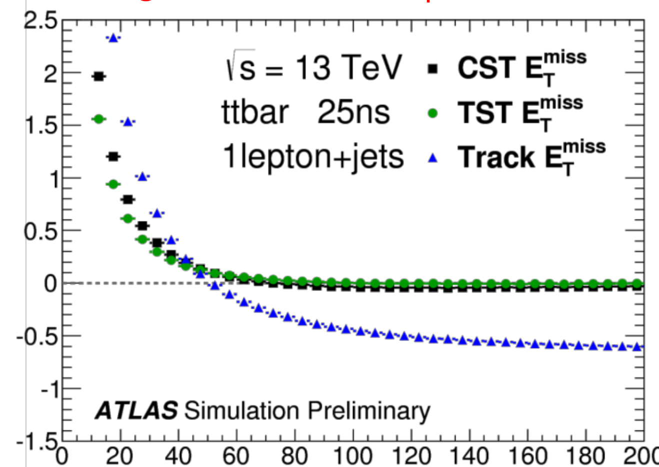
- In transverse plane, p_T(initial pp)=0; imbalance in final state ⇒ neutrino(s)
  - Need to measure 'everything else' from the hard-scatter, but avoid pileup

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss},\,e} + \cancel{E_{x(y)}^{\text{miss},\,\gamma}} + \cancel{E_{x(y)}^{\text{miss},\,\tau}} + E_{x(y)}^{\text{miss, jets}} + E_{x(y)}^{\text{miss},\,\mu} + E_{x(y)}^{\text{miss, soft}}$$

  - Reconstructed electrons, muons, hard jets (passing JVT selection)
  - Soft term is strongly polluted by pileup, use track-based soft term (TST) in run-2
    - Better than run-1 calo-based soft term (CST) – degrades at high sum(E_T) / high <μ>
    - Better than pure track-based E_T^{miss} –lacks neutral particles and jets with $|\eta|>2.5$
  - E_T^{miss} resolution from RMS in $Z\rightarrow\mu\mu$(+jets) with little true E_T^{miss}, and $t\bar{T}\rightarrow l\nu$+jets
  - Linearity check looking at relative bias in $t\bar{T}\rightarrow l\nu$+jets
    - Positive bias at low E_T^{miss} true (cannot measure –ve), drops for significant true E_T^{miss}

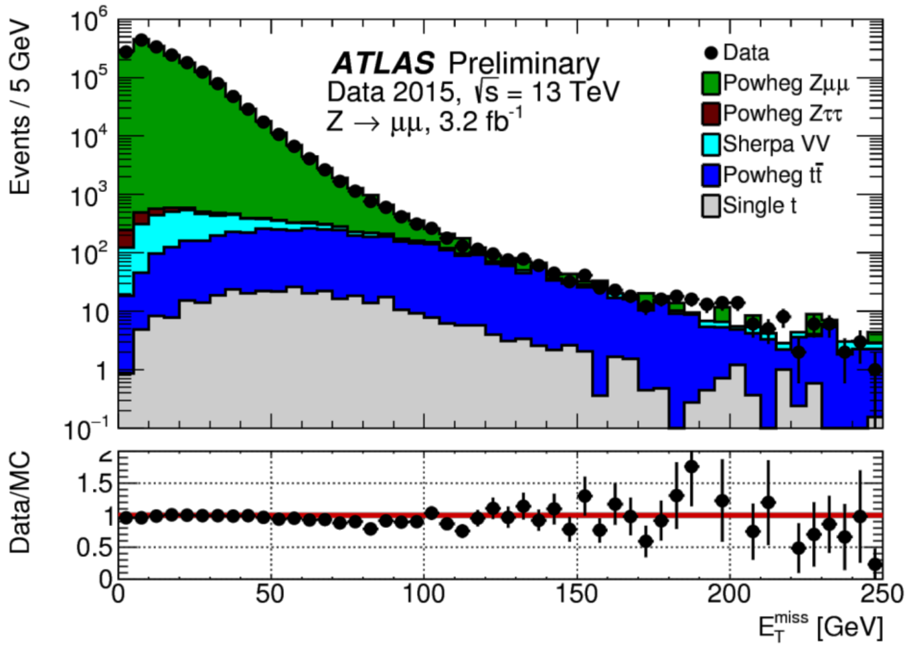


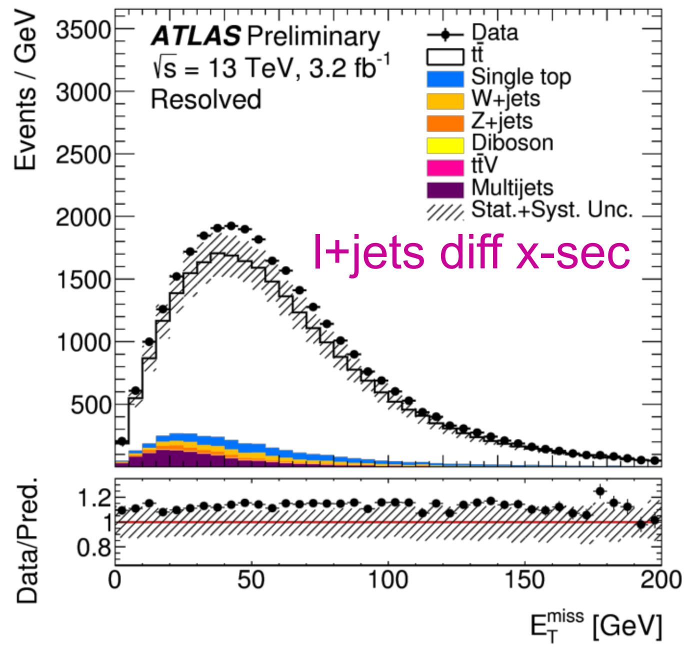




Missing transverse energy in action

- E_T^{miss} performance checked in a wide variety of samples
  - Component by component (systematic uncertantnies on each), and combined
  - Syst. on components (jets, e, μ) treated coherently between object and E_T^{miss}





- Cuts on E_T^{miss} (and W transverse mass) used l+jets event selections
  - Same-flavour dilepton selections (ee,$\mu\mu$) also usually cut on E_T^{miss}
- In kinematic fits, use W-mass constraint to estimate neutrino z-component
  - Dilepton events have two neutrinos – additional assumptions needed

Summary

- Top physics relies on many of the physics objects ATLAS can reconstruct
  - Electrons and muons for clean event signatures, straightforward triggering and precision measurements
  - Jets for reconstructing the complete final state, measuring kinematic properties and looking at jet activity associated to the top quark production
  - b-tagging to enhance event purity, identify the top decay jets
  - E_T^{miss} to aid in event selection, partially reconstruct the neutrino(s)
  - Complemented by use of taus (and photons) for specialist measurements
- All objects working well at run-2, object quality and calibration approaching run-1 values, despite harsher conditions (LHC energy, 25ns spacing, pileup)
  - Largest detector-related uncertainties typically coming from jets, especially when full event reconstruction is required
  - Lepton uncertainties typically smaller, thanks to precise calibration with $Z \rightarrow ll$ decays, though care needed to translate these results to top environment
  - Jet substructure techniques starting to bear fruit for boosted topologies
- Looking forward to more data and more top quarks …

References

- Overall: https://twiki.cern.ch/twiki/bin/view/AtlasPublic
- Electrons: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ElectronGammaPublicCollisionResults
  - ATLAS-CONF-2016-024, ATL-PHYS-PUB-2016-015, EPJC 74:2941
- Muons: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlots
  - EPJC 76:292, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonTriggerPublicResults
- Taus: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TauPublicCollisionPlots
- Jets/ ETMiss: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissPublicResults
  - arXiv:1510.08323, ATL-PHYS-PUB-2015-023, ATLAS-CONF-2016-040
- b-tagging: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FlavourTaggingPublicResultsCollisionData
  - ATL-PHYS-PUB-2015-022, ATL-PHYS-PUB-2016-12, ATL-PHYS-PUB-2015-039, ATLAS-CONF-2014-004,
- Tracking: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/InDetTrackingPerformanceApprovedPlots#Run_2
- Top and Z analyses
  - tT cross-section: EPJC 74:3109, PLB 761:136; Z cross-section: PLB 759:601
  - Top quark mass: EPJC 75:330, PLB 761:350, ATLAS-CONF-2016-064
  - tT gap fraction: EPJC 72:2043, JHEP 1609:074