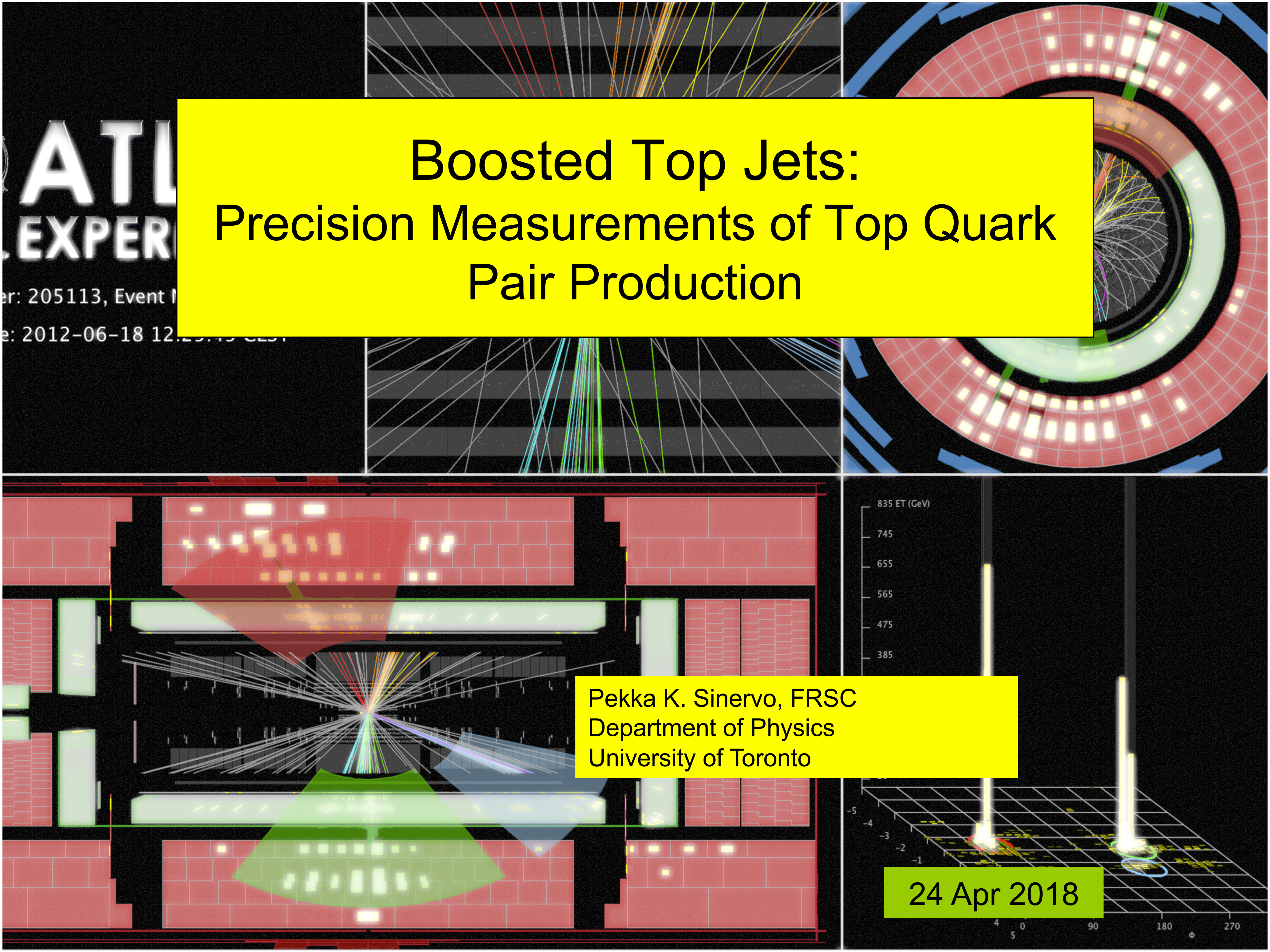
Boosted Top Jets:
Precision Measurements of Top Quark Pair Production

Pekka K. Sinervo, FRSC
Department of Physics
University of Toronto

24 Apr 2018

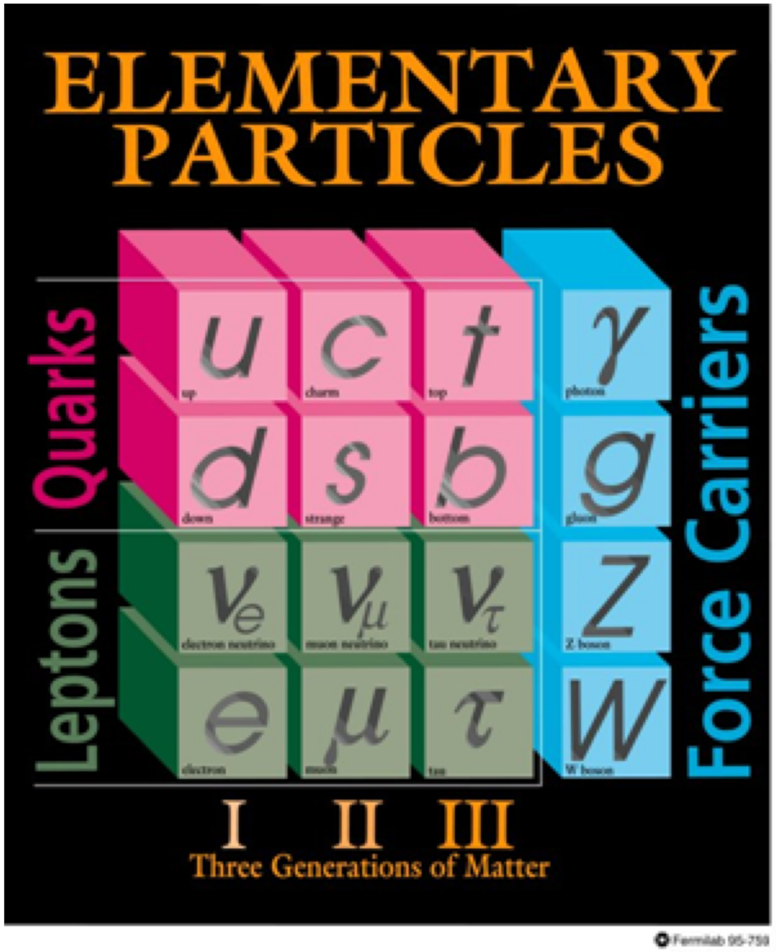


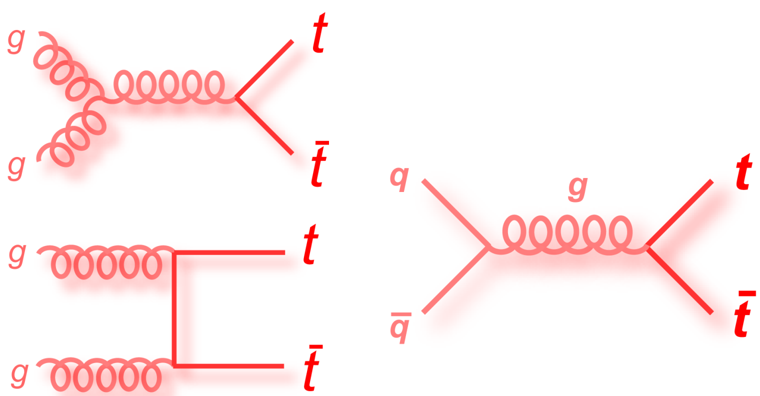
Outline

1. **Introduction to Boosted Top Quarks**
2. **QCD and Top Jets**
3. **Top and Bottom Tagging Algorithms**
4. **Differential Cross Section Measurements**
5. **Summary & Conclusions**

The Top Quark

- **Unique role**
  - Heaviest particle
  - In SM, strongest coupling to Higgs boson
  - A unique role in various extensions
- **QCD makes precise predictions for production**
  - NNLO predictions exist





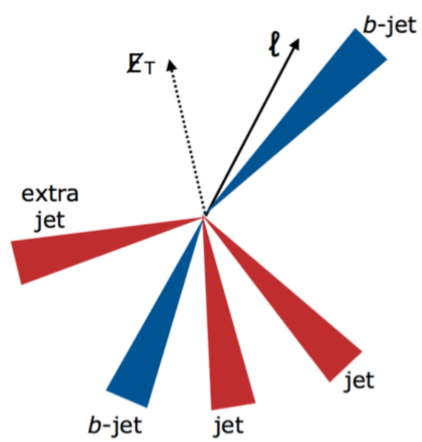
Top Quark is Special

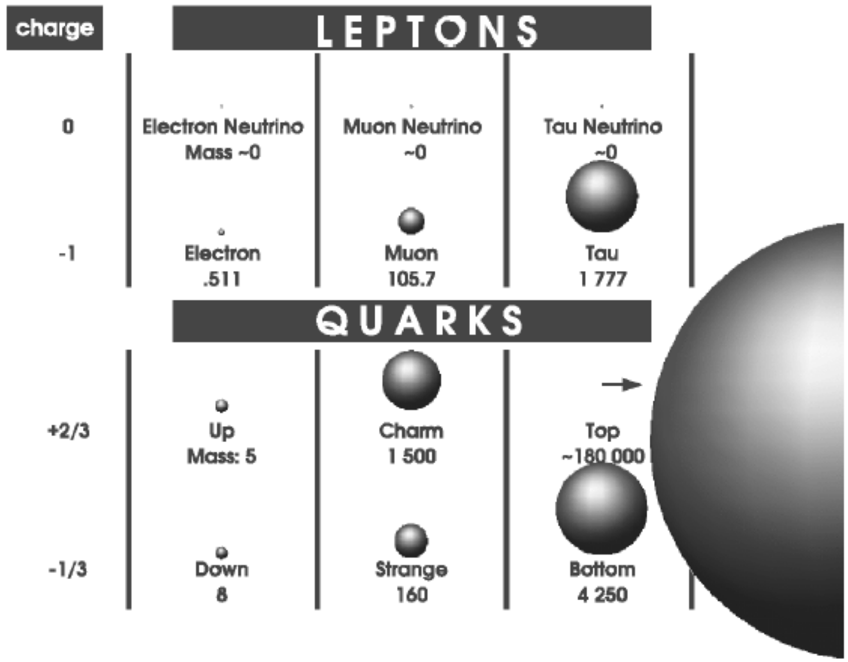
Properties

- Most massive fermion
- Decays before hadronization

Large mass also a challenge

- Decay products are energetic
- Difficult to reconstruct top quark decays in pair production





Incandela, Quadt, Wagner & Wicke, Prog.Part.Nucl.Phys.63:239-292,2009

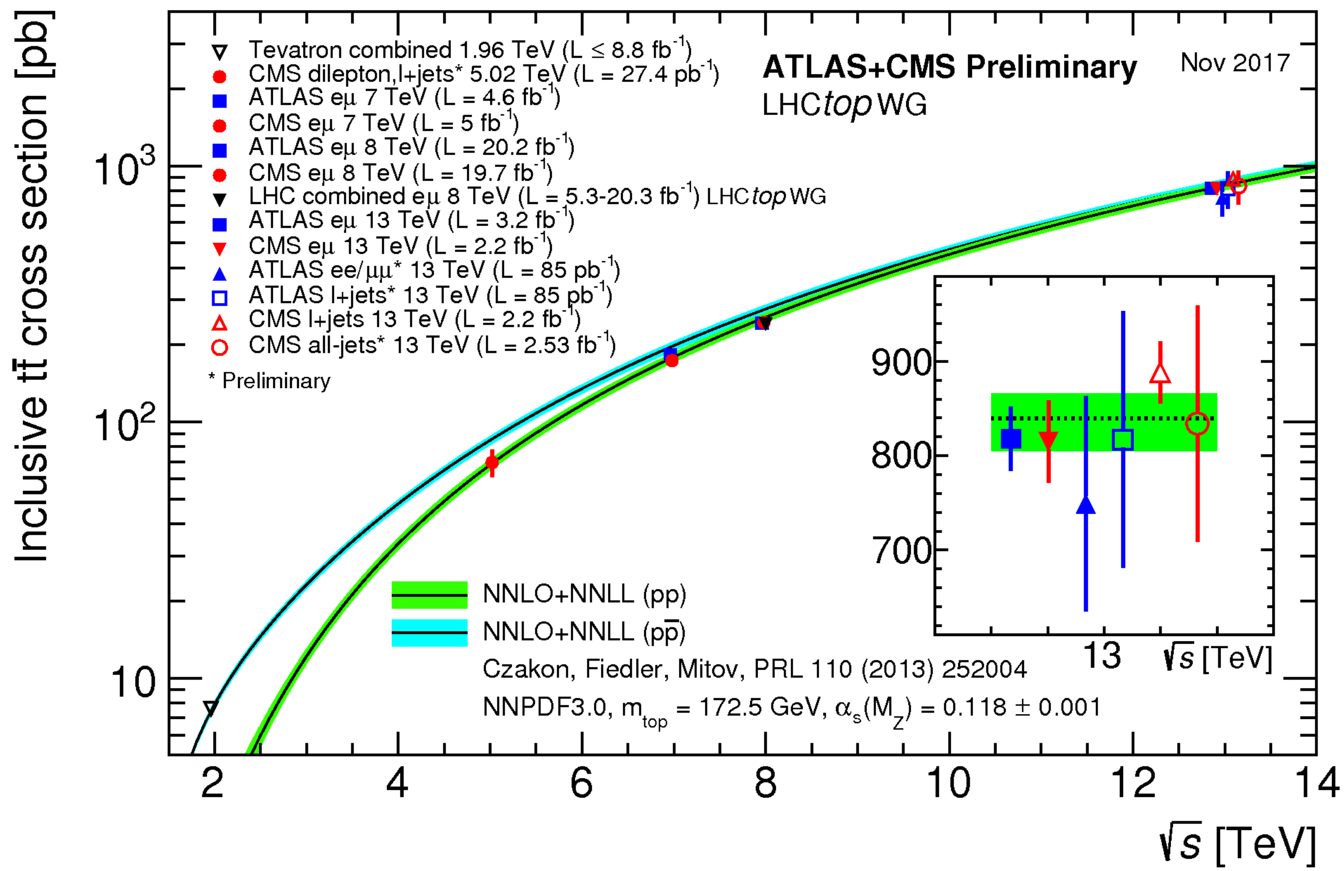
What Are Boosted Top Quarks?

Top quark pair production ubiquitous at LHC

- $\sigma = 832 \pm 51$ pb

LHC TOP WG

- At 10^{34} cm^{-2} s^{-1}, this gives a top quark every few seconds



Boosted top quarks are those with

$p_T \gtrsim 500$ GeV

- Rare – 0.1% of total

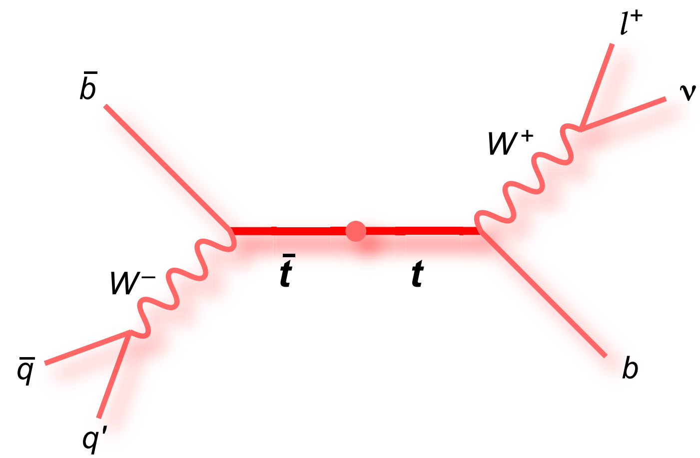
Top Quark Pair Production

Each top quark decays weakly

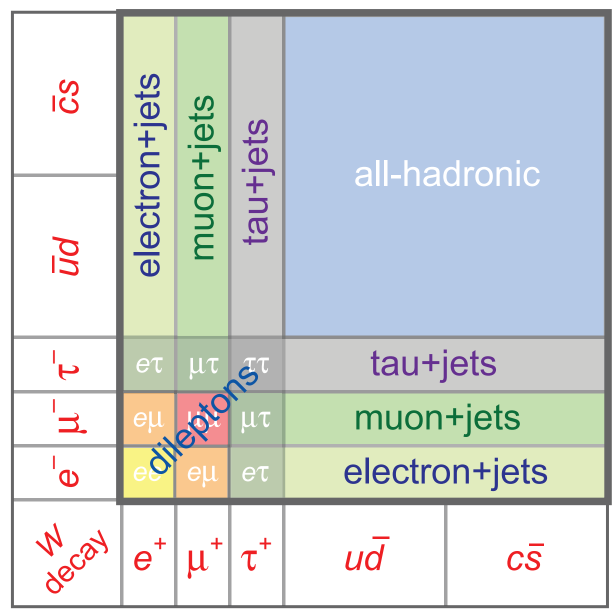
- 3 possible final states

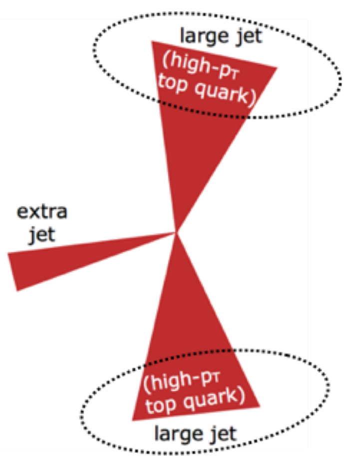
Focus on the "all-hadronic" mode

- 44% of all decays
- No undetected particles!
- At high p_T, top quark decay reconstructed as a "top jet"



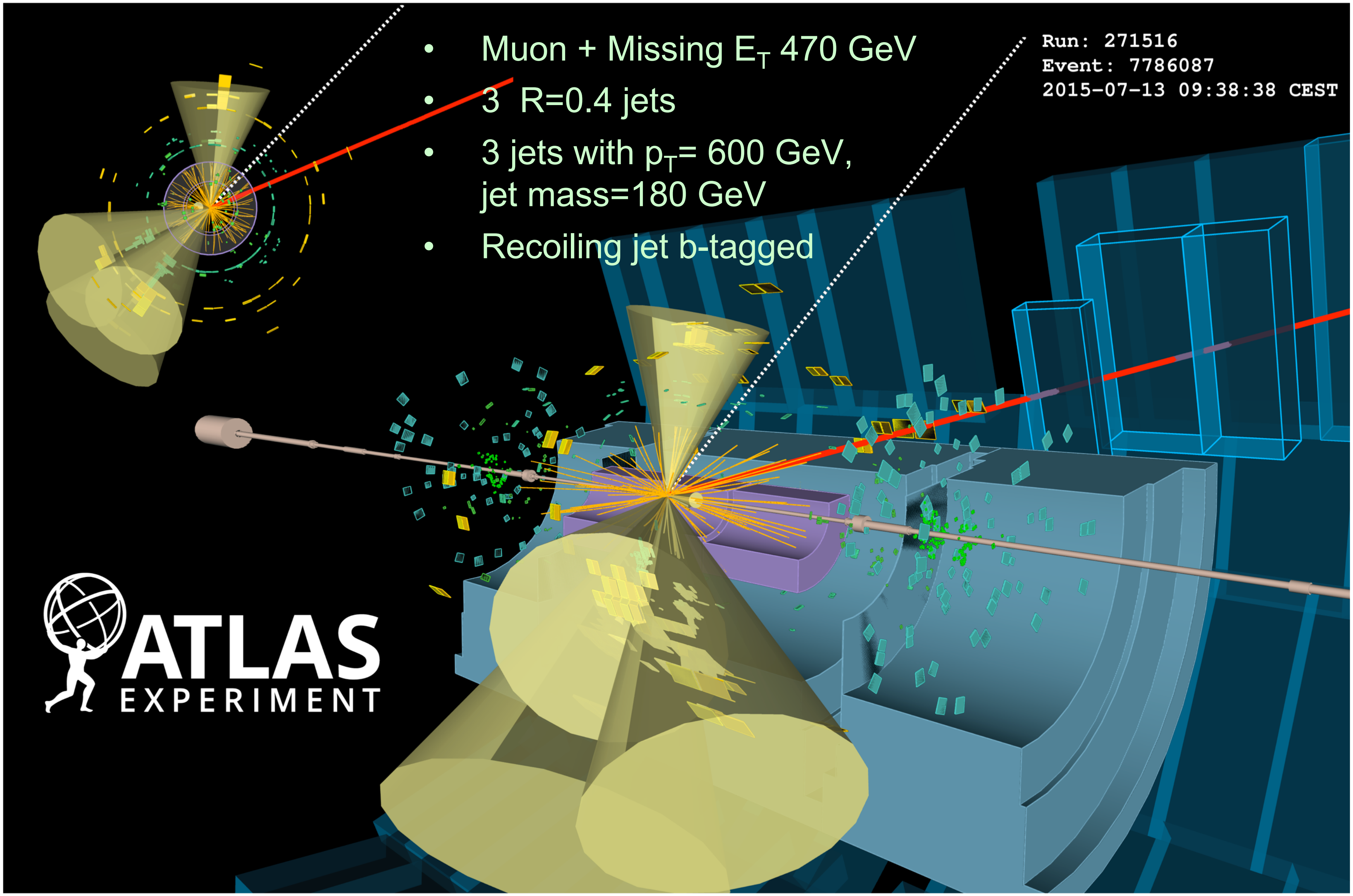
Top Pair Decay Channels





A Very high-P_T Top

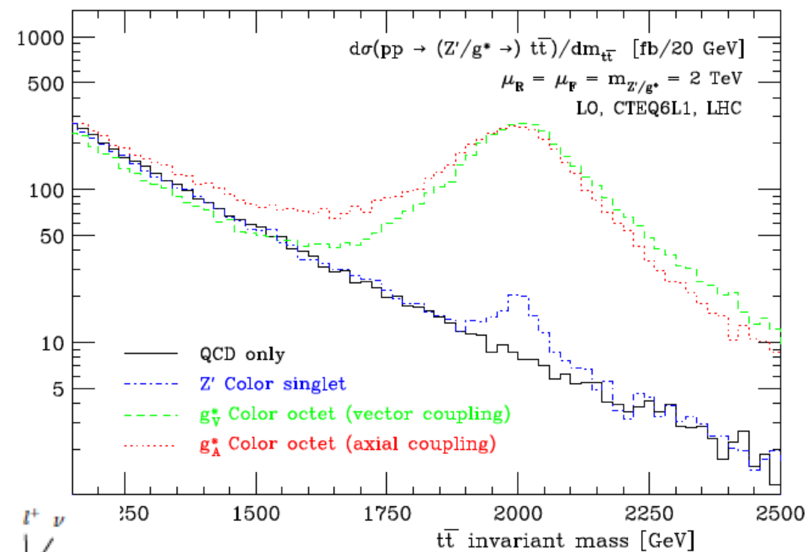
- Muon + Missing E_T 470 GeV
- 3 R=0.4 jets
- 3 jets with p_T= 600 GeV, jet mass=180 GeV
- Recoiling jet b-tagged



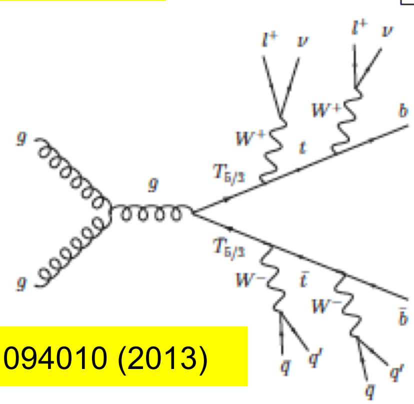
Why the Interest in Boosted Tops?

Top quarks play a special role in many models for new physics, eg.

- Couple to new force carriers
  - Leptophobic Z' preferentially decays to top quark pairs
    Rosner, PLB **387** (1996) 113
  - W' bosons could decay to t-b pair
- String-inspired resonances
  - Randall-Sundrum KK gluons/gravitons (g_{KK},G_{KK}) favourite "wide" resonance
    Agashe et al., PRD **77**, 015003 (2008) ;
    Lillie et al., JHEP **09** (2007) 074
- New phenomena
  - Vector-like top quark partners
  - Supersymmetric top partners



Fredrerix & Maltonii, JHEP **01** (2009) 047



Agular-Saavedra et al, PRD **88**, 094010 (2013)

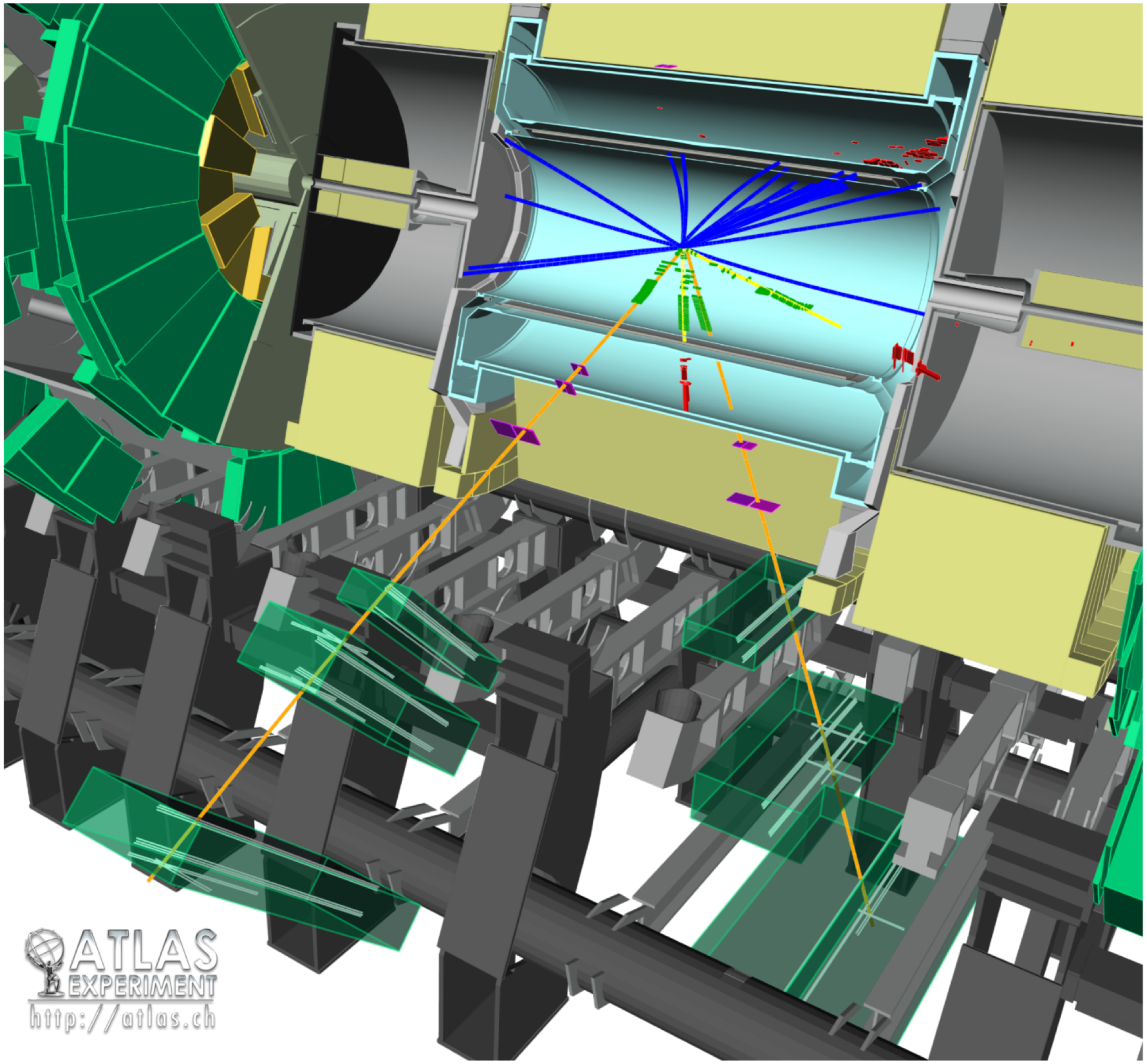
CERN Large Hadron Collider

LHC collides protons at 13 TeV

High luminosity

- 2800 bunches
- 1.2×10^{11} protons per bunch
- Bunch crossings every 25 ns
- $L > 10^{34}\ cm^{-2}\ s^{-1}$

The ATLAS Detector



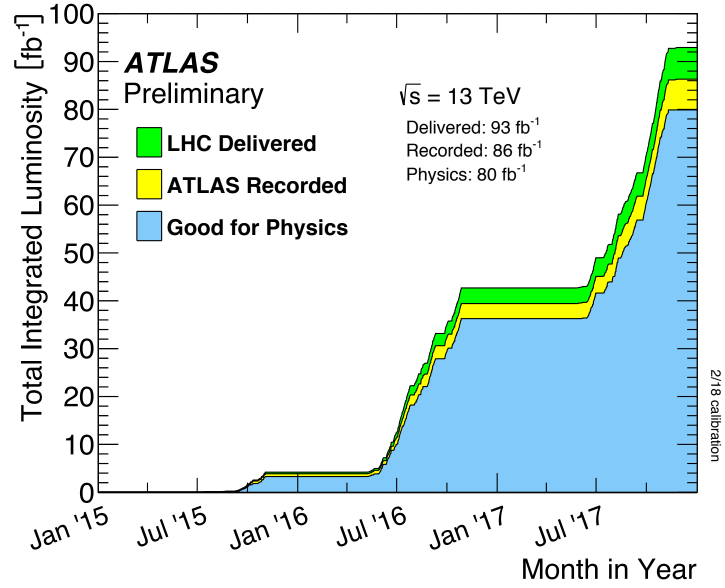
LHC Data Taking

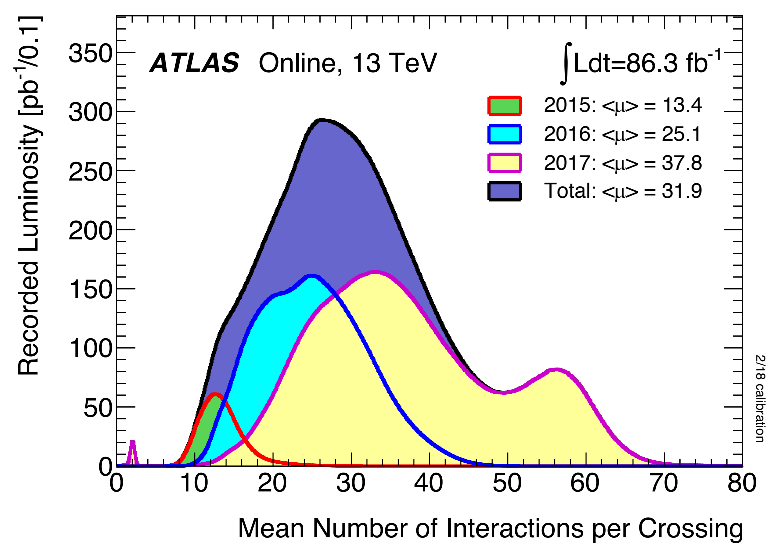
LHC has run well at 13 TeV

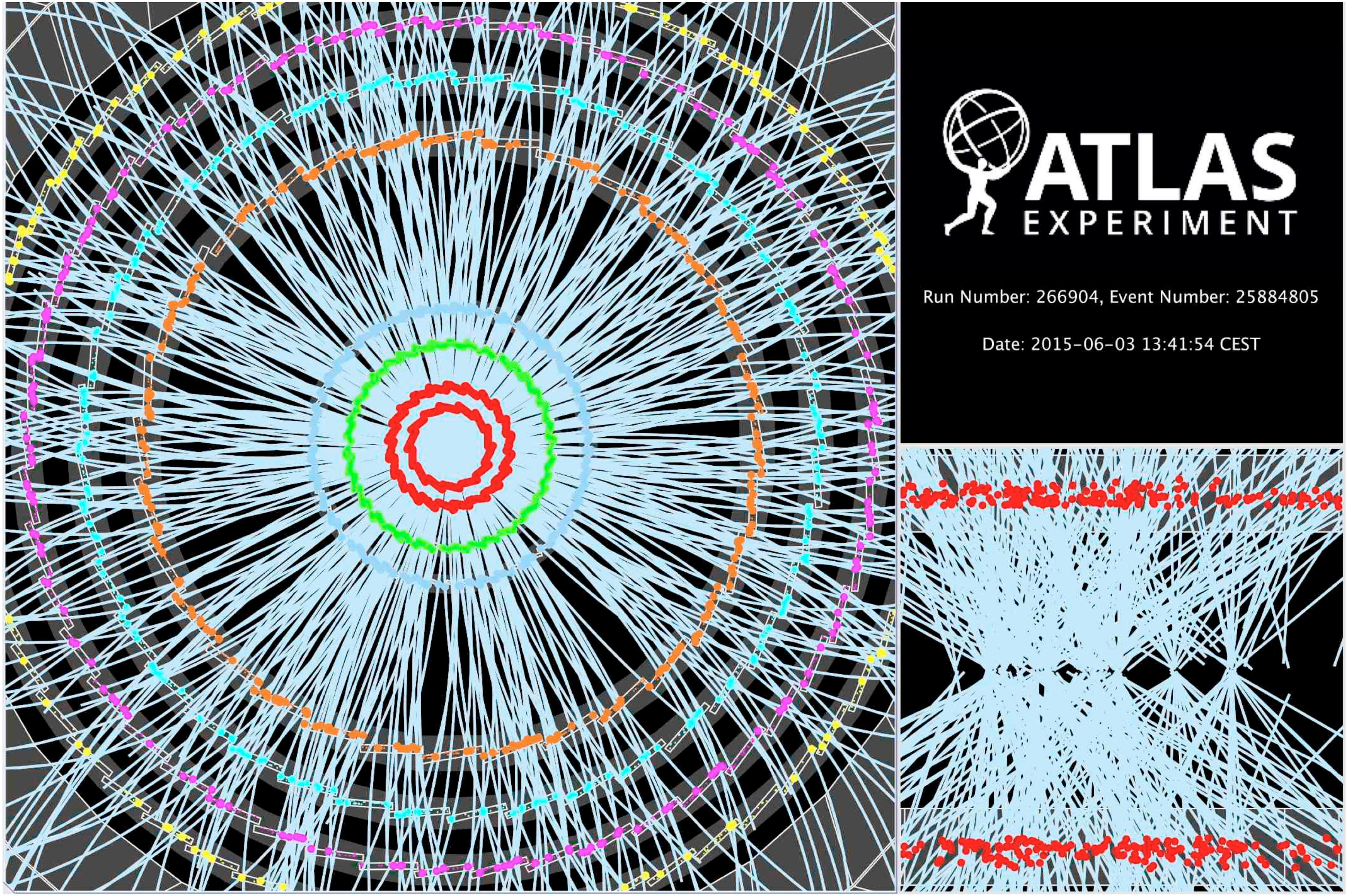
13 TeV data:

- 3.9 fb^{-1} in 2015
- 32 fb^{-1} in 2016
- 35 fb^{-1} in 2017

Collision "pile-up" continued to increase





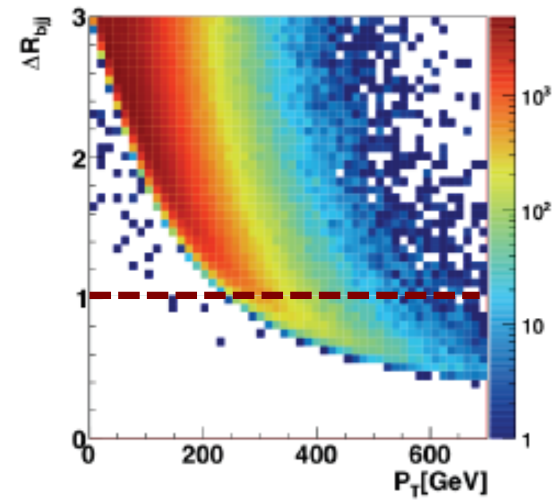


We believe there were 17 collisions…

High p_T Top Jets

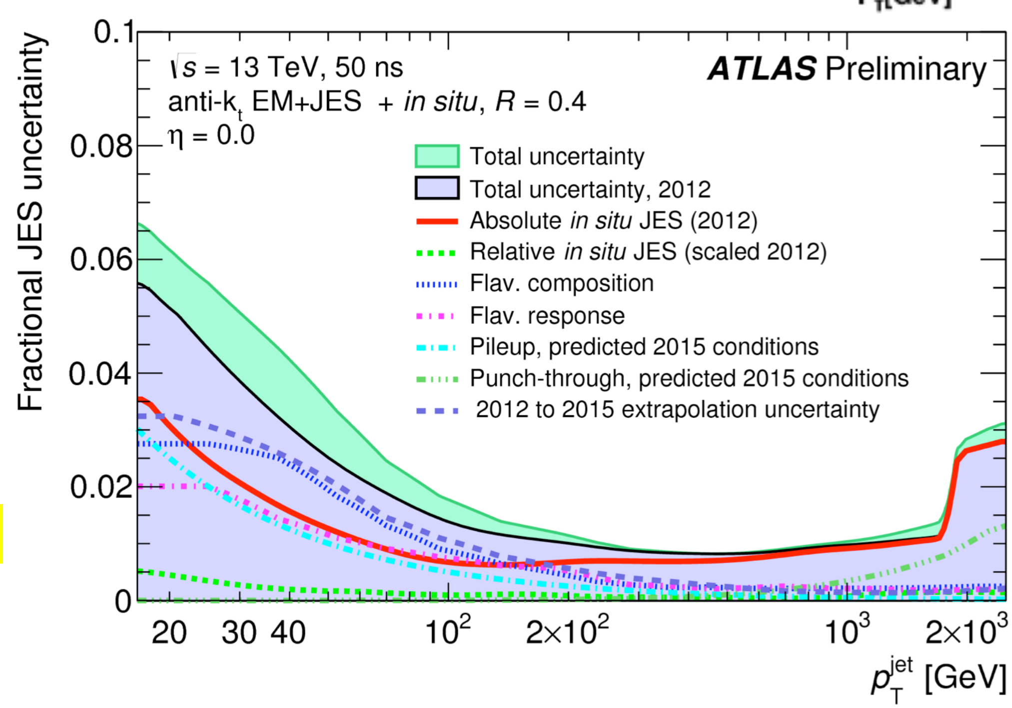
High p_T top decays seen as single jet

- Use Anti-kt jet algorithm
  - R=1.0 to capture top decay products
  - But sensitive to pile-up



- Calibrate energy and mass scales using *in situ* techniques

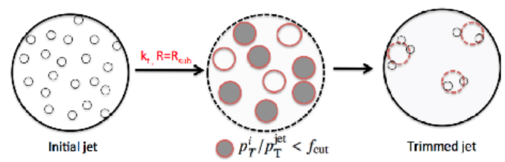
ATLAS-PUB-2015-015



Jet Trimming

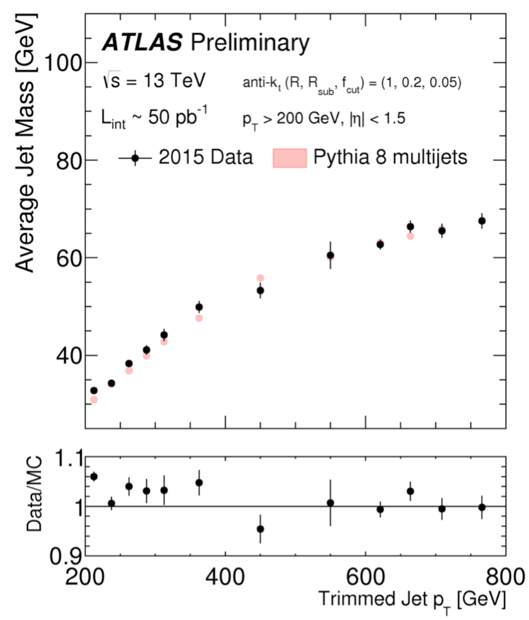
Pile-up is removed using "trimming"

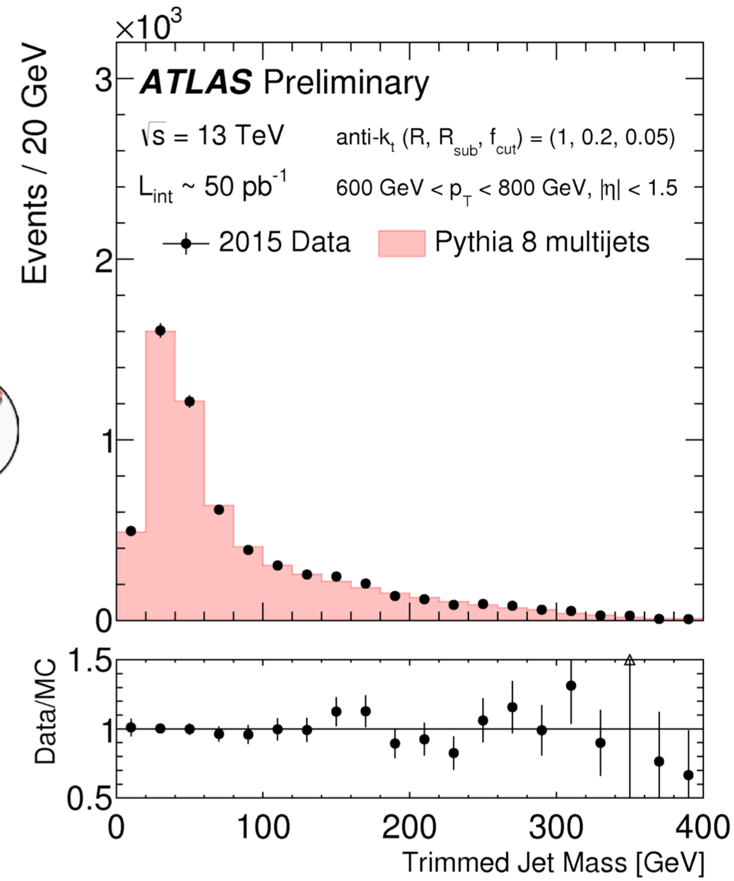
- Anti-kT cluster with R=1.0 – $p_T^{R1.0}$
- Anti-kT cluster constituents into R=0.2 "subjets"
- Keep subjets with pT > 0.05 $p_T^{R1.0}$
- Recombine and re-calibrate



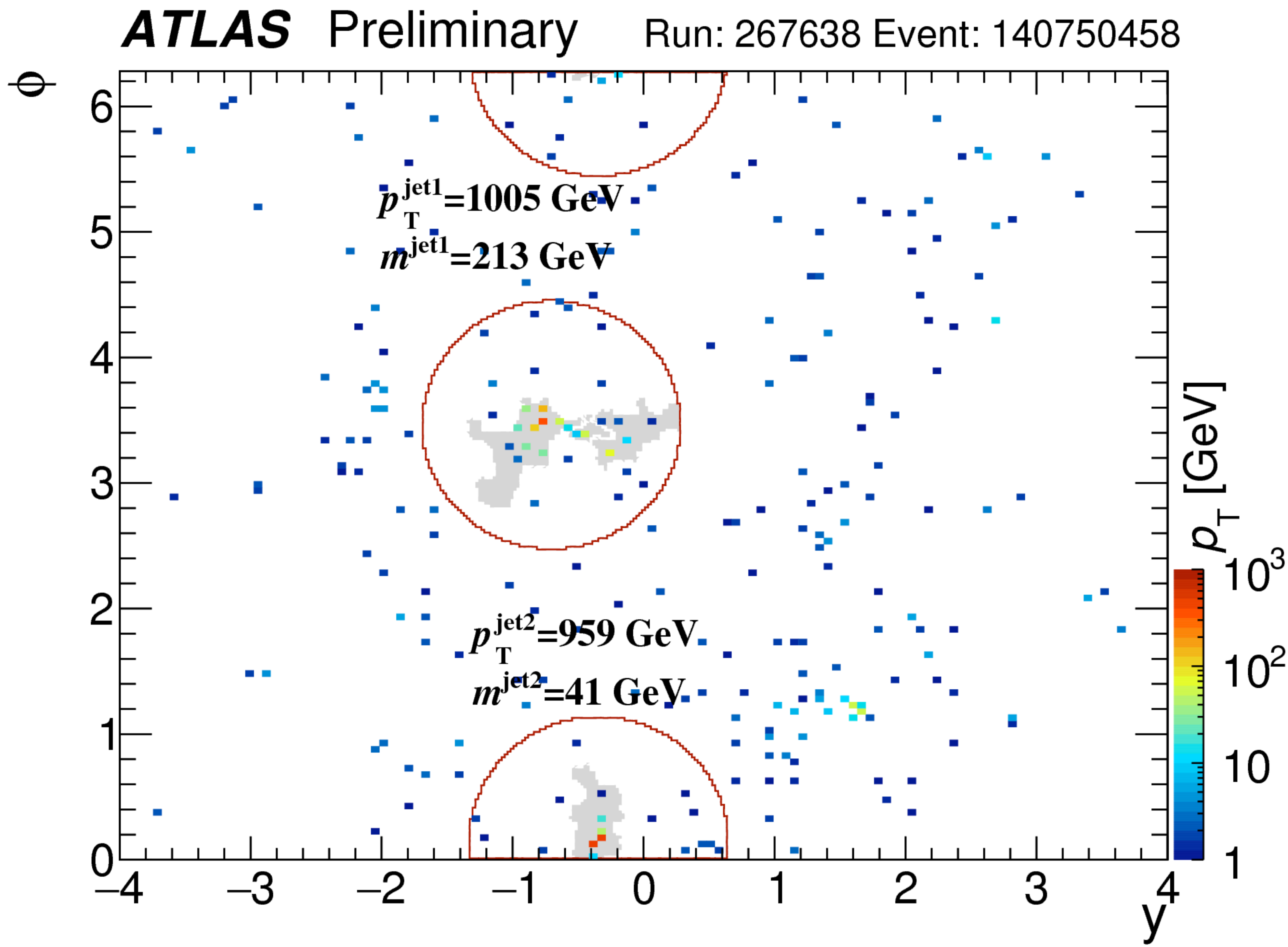
Takes care of pile-up

- Also "suppresses" Sudakov peak
- Rises slowly with jet p_T





ATLAS-CONF-2015-035


ATLAS Preliminary
Run: 267638 Event: 140750458
p_T^{jet1}=1005 GeV
m^{jet1}=213 GeV
p_T^{jet2}=959 GeV
m^{jet2}=41 GeV
φ
y
p_T [GeV]

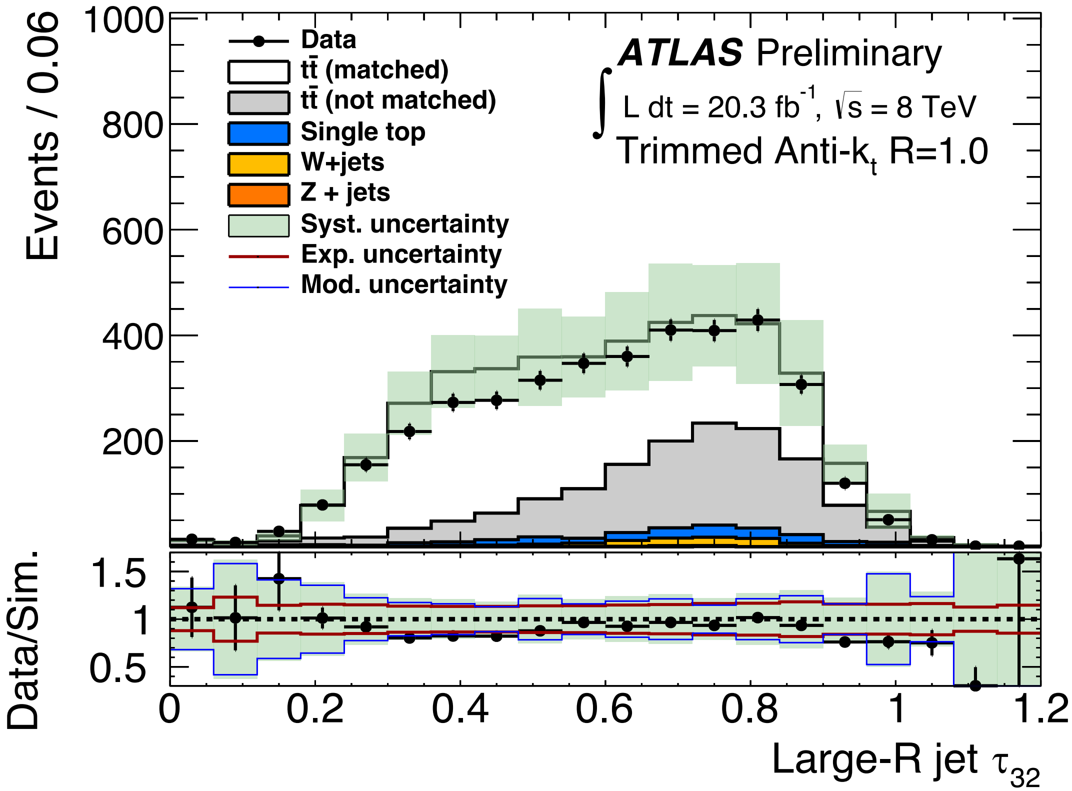
Jet Mass Isn't Everything

Top quark decays are 3-prong

- Light quark and gluon jets with high mass largely single gluon emission

Many strategies considered

- Eight algorithms compared in ATLAS-CONF-2015-036
- Taken a simple approach for "top-tagger"
  - N-subjettiness measure τ_{32}
  - Jet mass cut ($122 < m_{jet} < 222$ GeV)



Thaler & van Tilburg, JHEP 03 (2011) 015; JHEP 02 (2012) 093

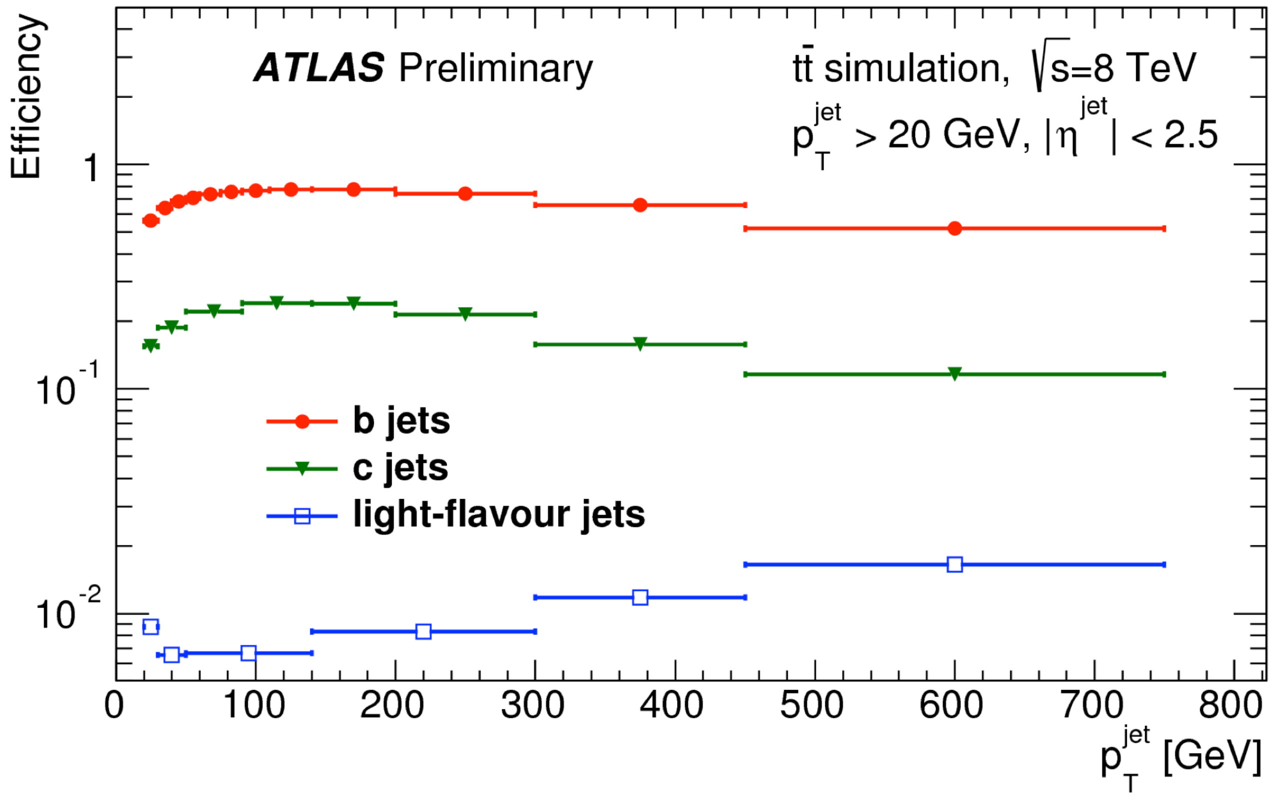
B Tagging Algorithms

ATLAS uses a multivariate algorithm to tag "b-jets"

- Combination of tracking, vertex and kinematic information
- Usual operating point of 70% efficiency, <1 % mistag rate

- Require R=0.4 b-tagged jet associated with top quark jet

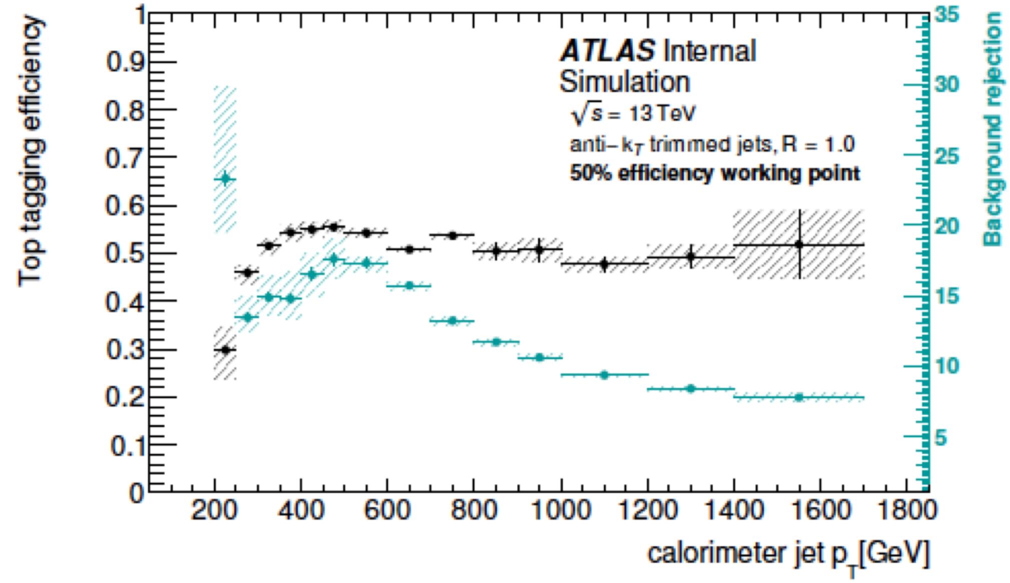
Essential element of tagging strategy

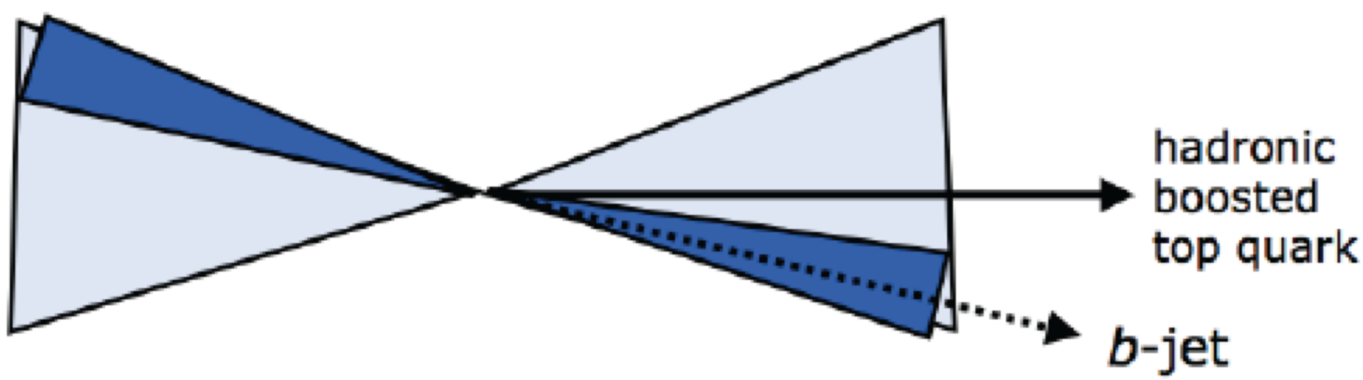


Tagging Pairs of Top Jets

Put together top-tagging and b-tagging

- Require two R=1.0 jets
  - $p_{T1} > 500$ GeV and $p_{T2} > 350$ GeV
  - Require both are top-tagged
  - Require both have R=0.4 subjet that is b-tagged
- Reject events with at least one electron or muon candidate





Estimating Backgrounds

Strategy

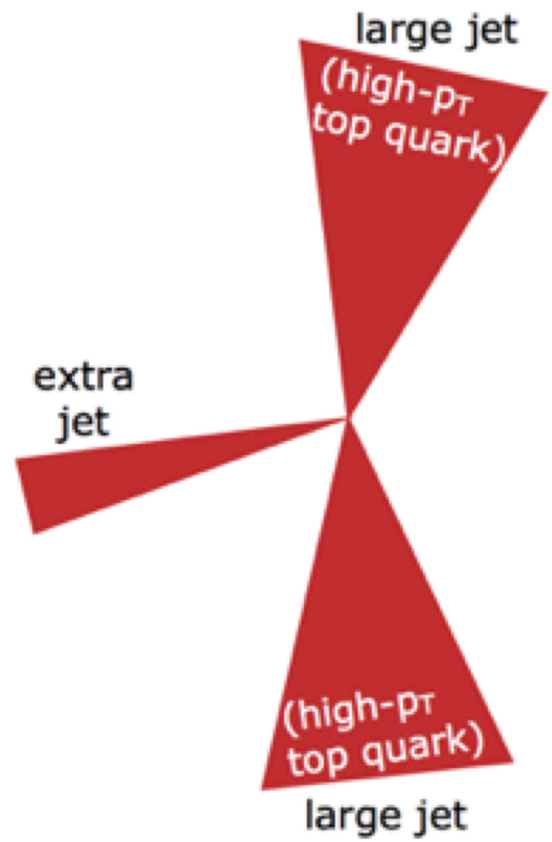
- Use MC for backgrounds with at least one real top quark jet:

  $t\bar{t}$ (non-all-hadronic)

  Single-top-quark

  $t\bar{t}$+$W/Z/H$

- Use data to estimate "multijet" background



2nd large-R jet				
1t1b	J (7.6%)	K (21%)	L (42%)	S
0t1b	B (2.2%)	D (5.8%)	H (13%)	N (47%)
1t0b	E (0.7%)	F (2.4%)	G (6.4%)	M (30%)
0t0b	A (0.2%)	C (0.8%)	I (2.2%)	O (11%)
	0t0b	1t0b	0t1b	1t1b

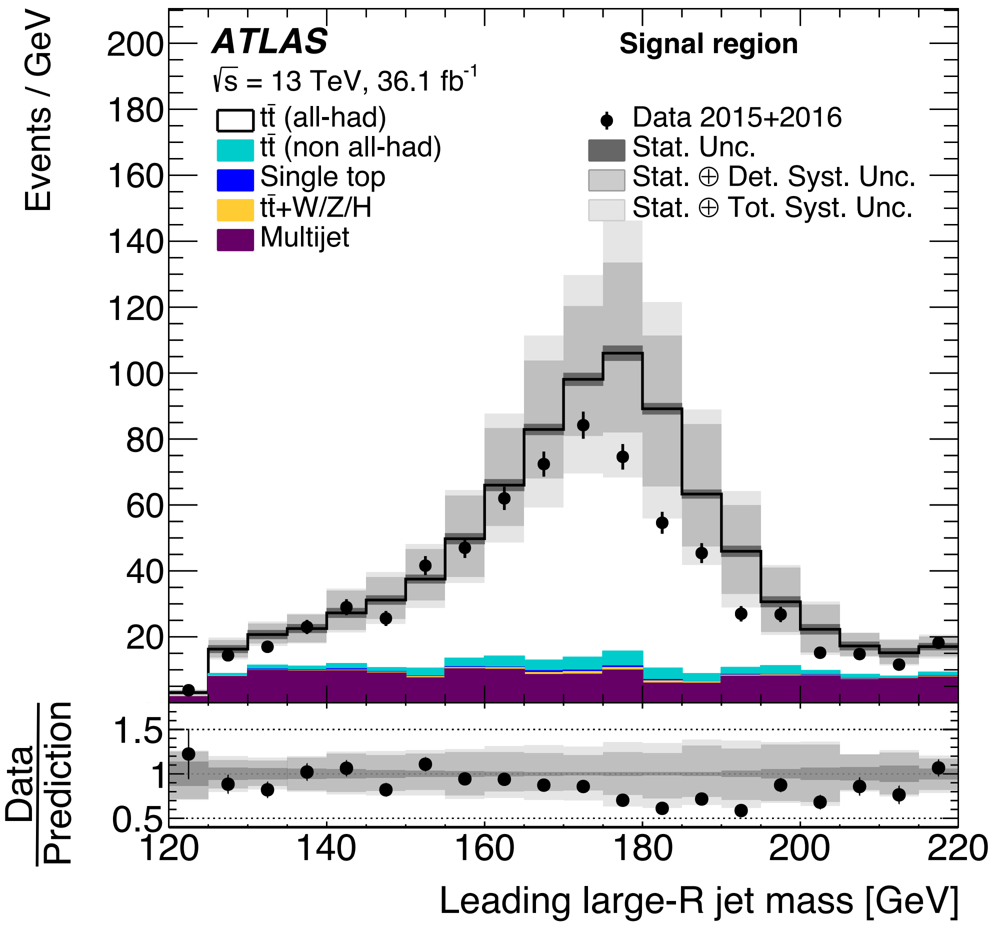
Leading large-R jet

What the Signal Looks Like

Jet mass is best diagnostic

- Clear top jet peak
- Background is flat

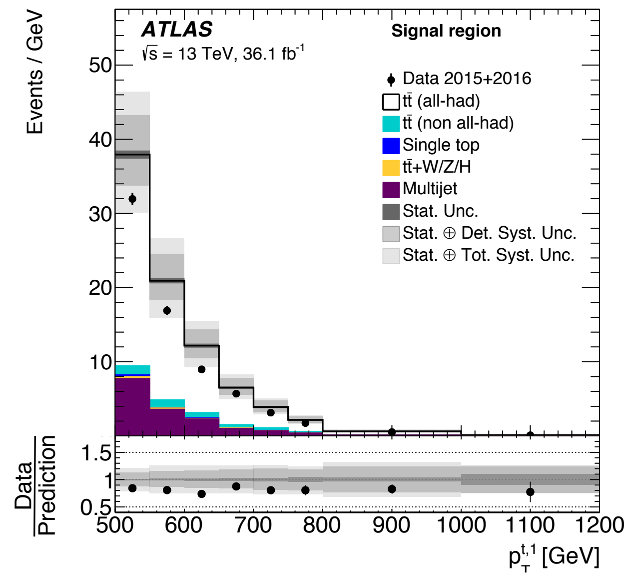
$t\bar{t}$ (all-hadronic)	3250 ± 470
$t\bar{t}$ (non-all-hadronic)	200 ± 40
Single-top-quark	24 ± 12
$t\bar{t}$+$W/Z/H$	33 ± 10
Multijet events	810 ± 50
Prediction	4320 ± 530
Data ($36.1\,\text{fb}^{-1}$)	3541



Kinematics of Events

Event properties

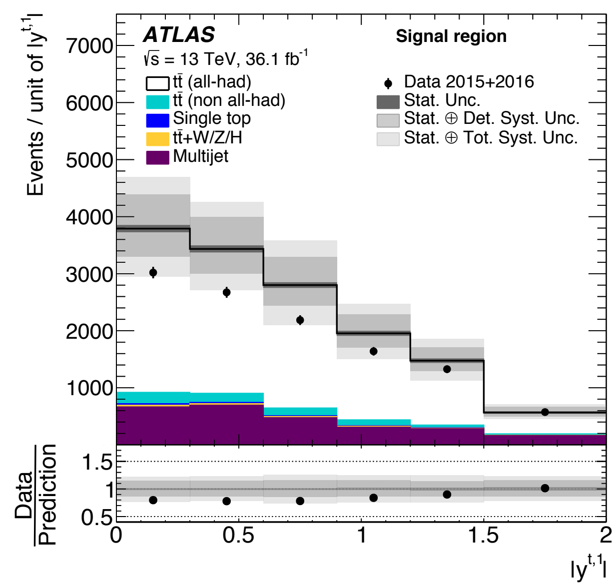
- Rapidly falling vs p_T
- Back-to-back pair production
- Centrally produced

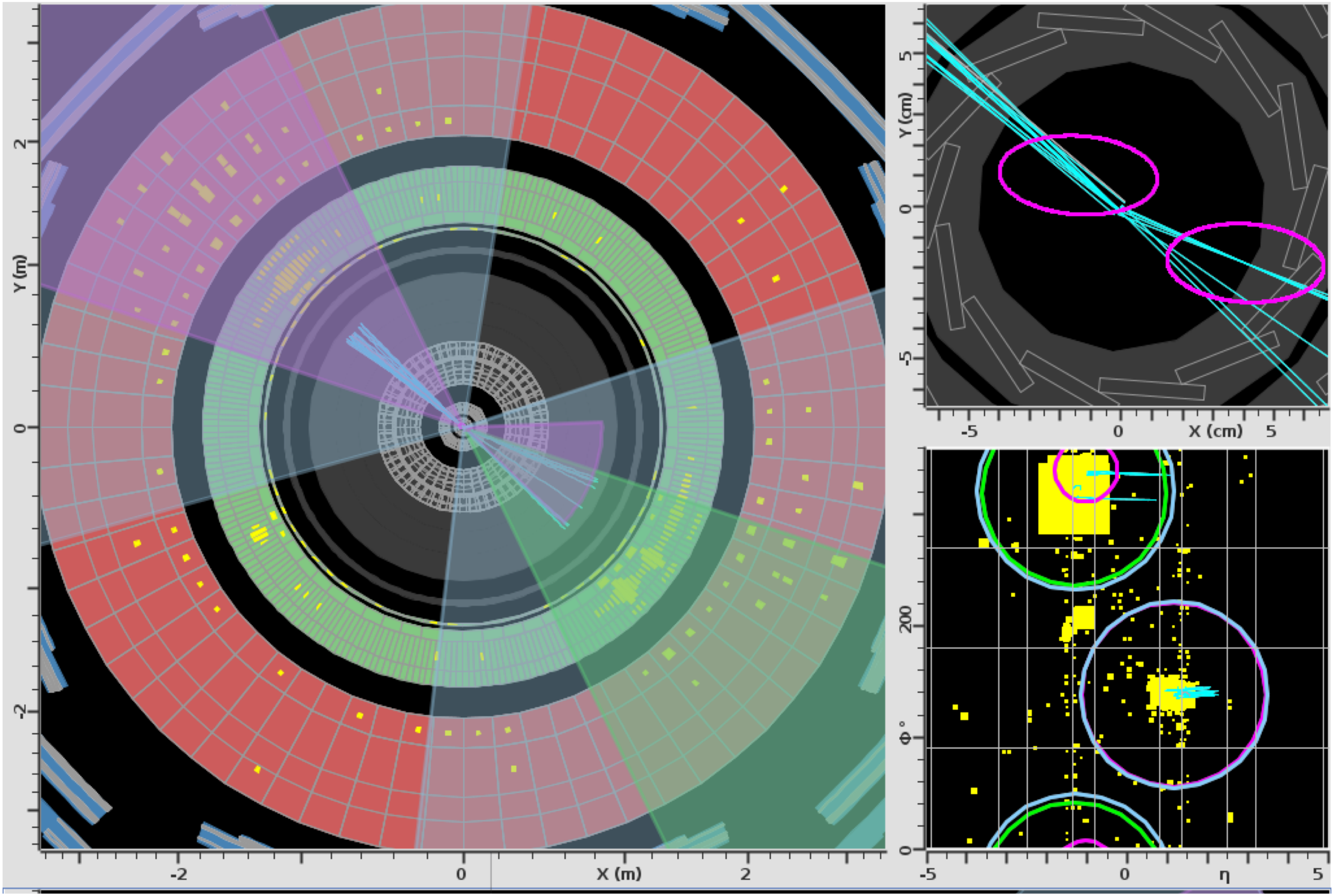


Affected by acceptance & efficiency

- Correct by ”unfolding” distributions
- Measure normalized differential cross sections vs. 13 variables

$$\frac{1}{\sigma^{fid}}\frac{d\sigma^{fid}}{dX_i}$$



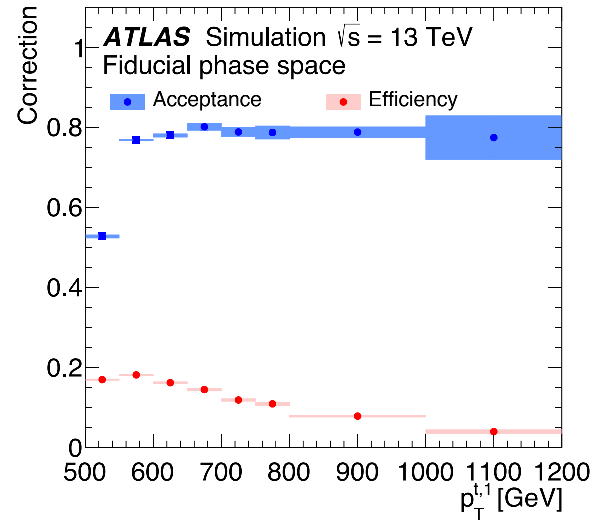


Y (m)
2
0
-2
-2
0
X (m)
2
Y (cm)
5
0
-5
-5
0
X (cm)
5
200
Φ°
0
-5
0
η
5

Unfolding to Particle Level

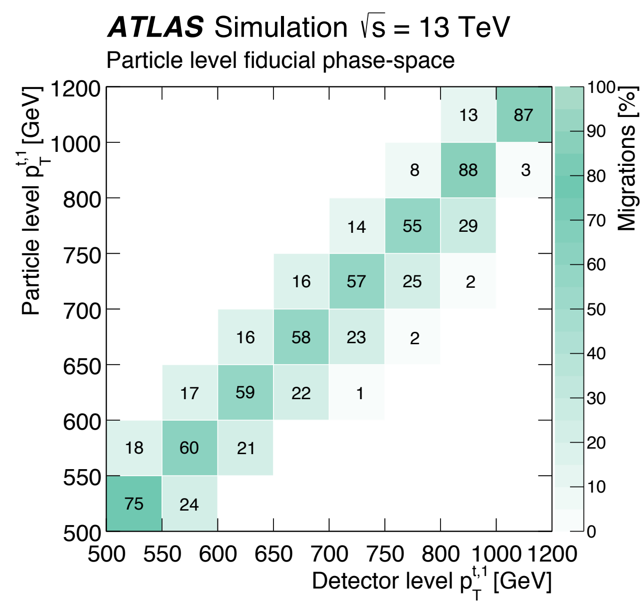
Define a "fiducial phase space":

- Large-R jet $p_T^{t,1}$> 500 GeV
- Large-R jet $p_T^{t,2}$> 350 GeV
- Each associated with b-tag



Use a Bayesian unfolding procedure to correct for acceptance & efficiency

$$\frac{d\sigma^{\text{fid}}}{dX^i} \equiv \frac{1}{\int \mathcal{L}\,dt \cdot \Delta X^i} \cdot \frac{1}{\epsilon_{\text{eff}}^i} \cdot \sum_j \mathcal{M}_{ij}^{-1} \cdot f_{\text{acc}}^j \cdot \left(N_{\text{reco}}^j - N_{\text{bg}}^j\right)$$

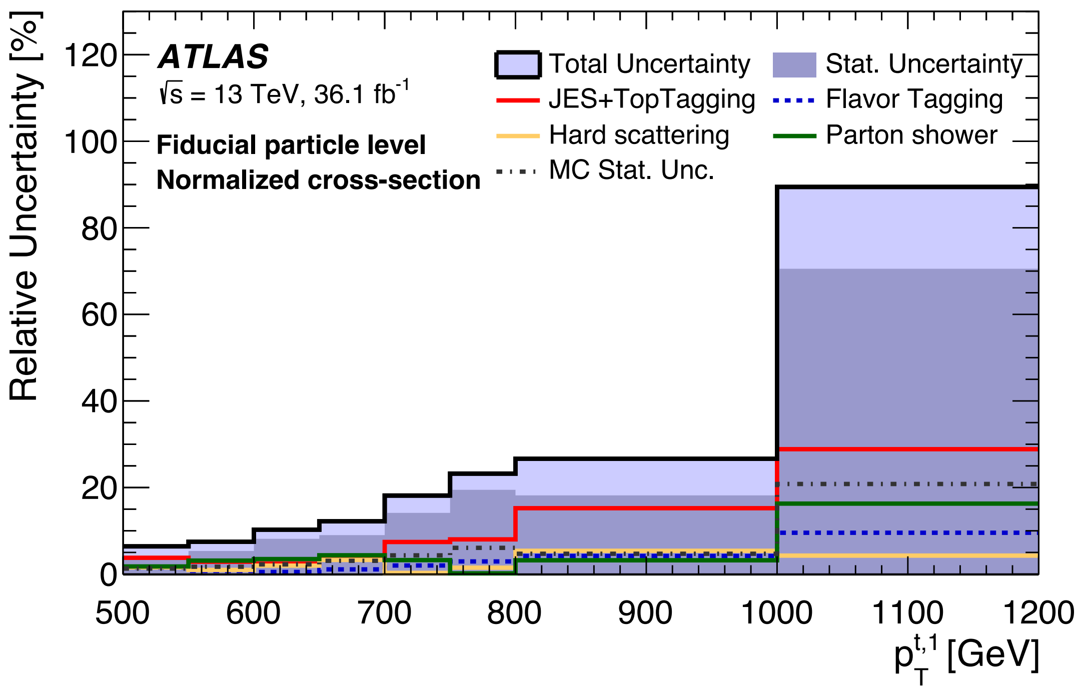


Systematic Uncertainties

Estimate systematic uncertainties

- Based on studies of data samples
- Modelling uncertainties come from MC comparisons
- Incorporated using Bayesian priors

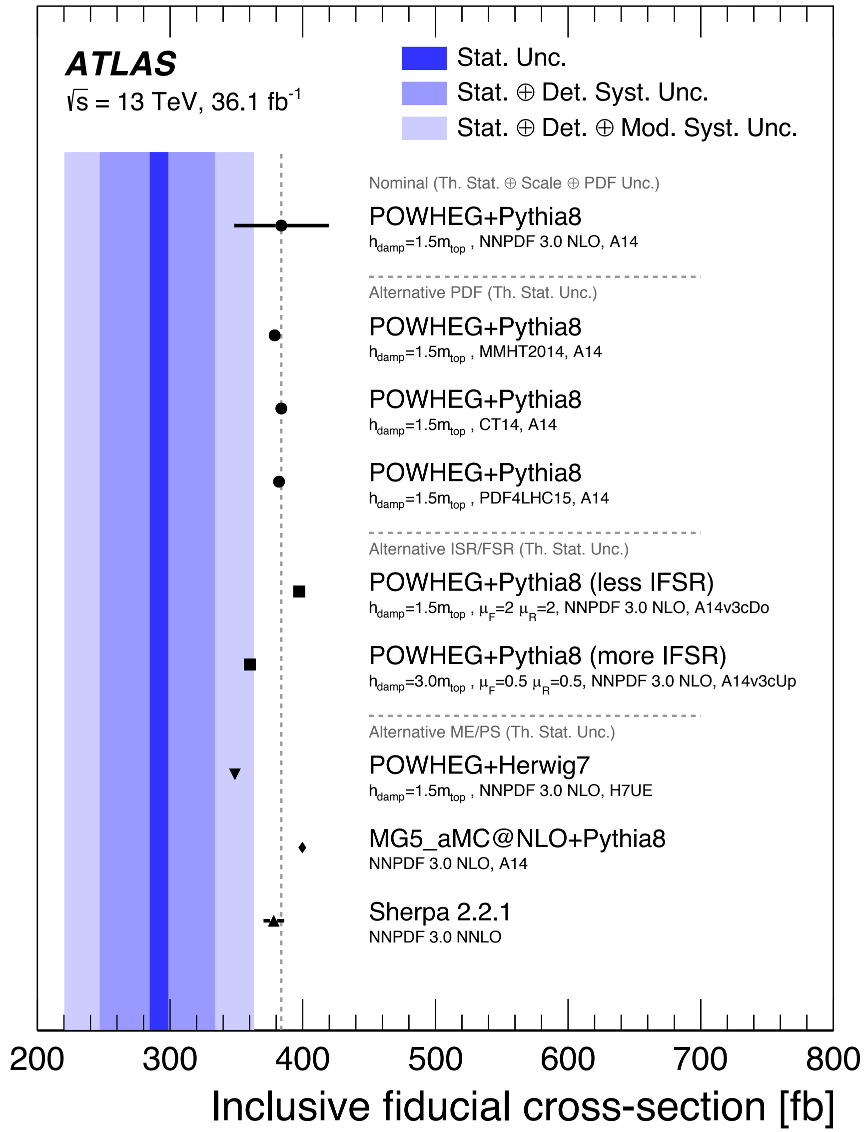
Source	Percentage
Large-R jet energy scale	5.9
Large-R jet mass calibration	1.4
Large-R jet top-tagging	12
Small-R jets	0.3
Pileup	0.6
Flavor tagging	8.3
Background	0.9
Luminosity	2.0
Monte Carlo statistical uncertainty	0.9
Alternative hard-scattering model	11
Alternative parton-shower model	14
ISR/FSR + scale	1.1
Total systematic uncertainty	24
Data statistical uncertainty	2.3
Total uncertainty	24



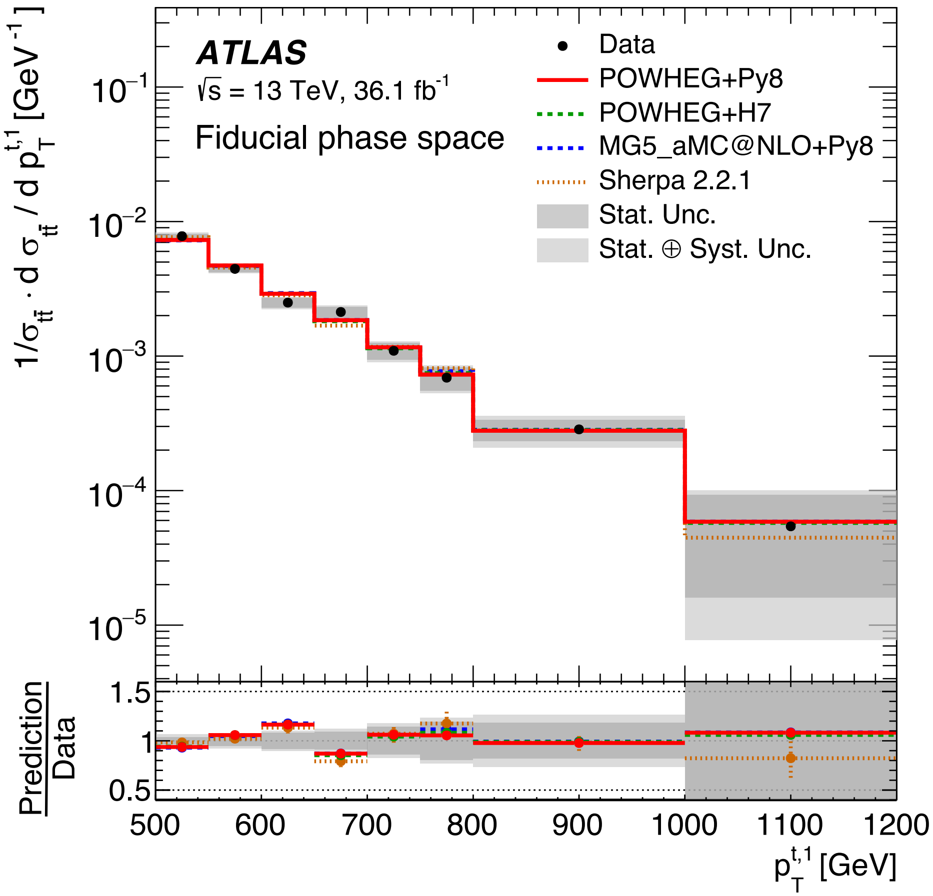
Total Cross Section Measurement

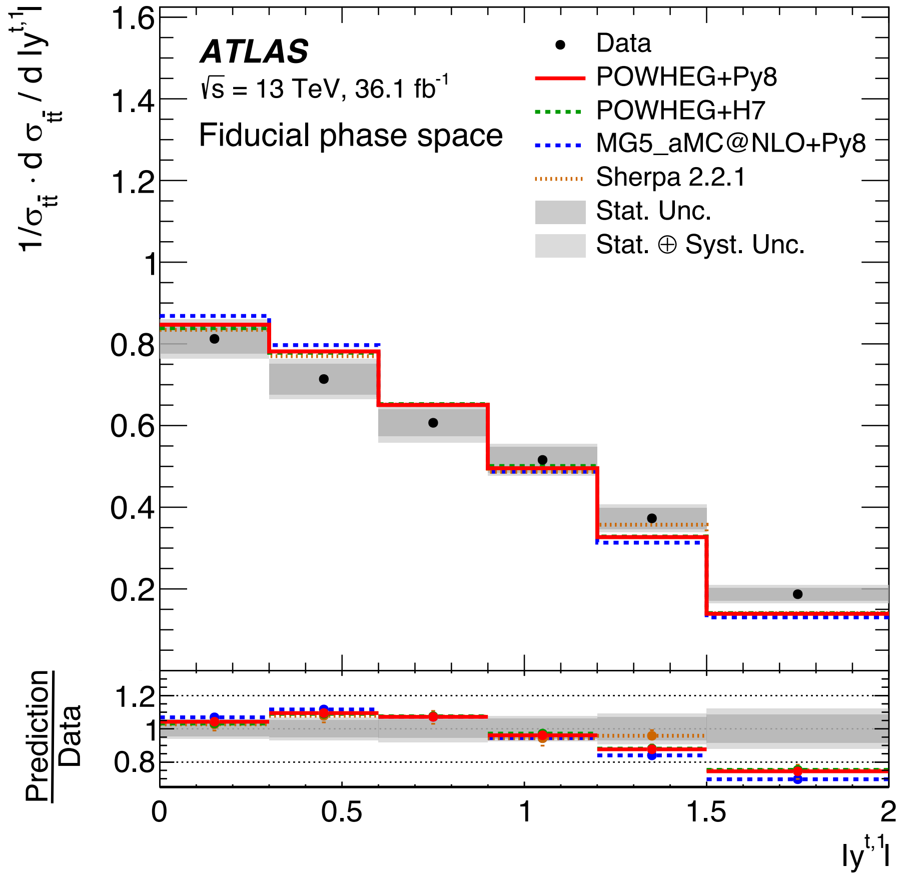
Compare with different matrix-element calculations

- POWHEG
  - Pythia v8 parton shower + hadronization
  - Herwig v7 parton shower + hadronization
- MG5_aMC@NLO
  - Pythia v8 parton shower + hadronization
  - Herwig v7 parton shower + hadronization
- SHERPA 2.2.1

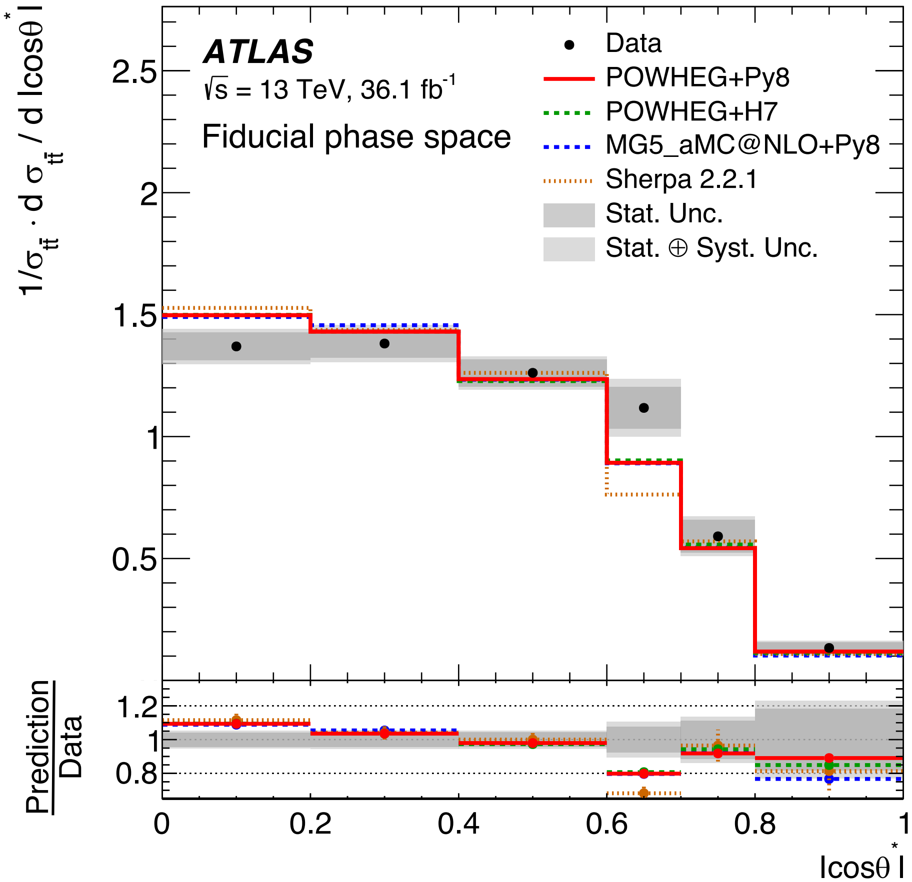


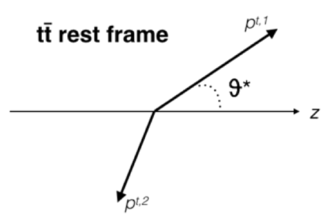
Unfolding Results: Leading Top Jet

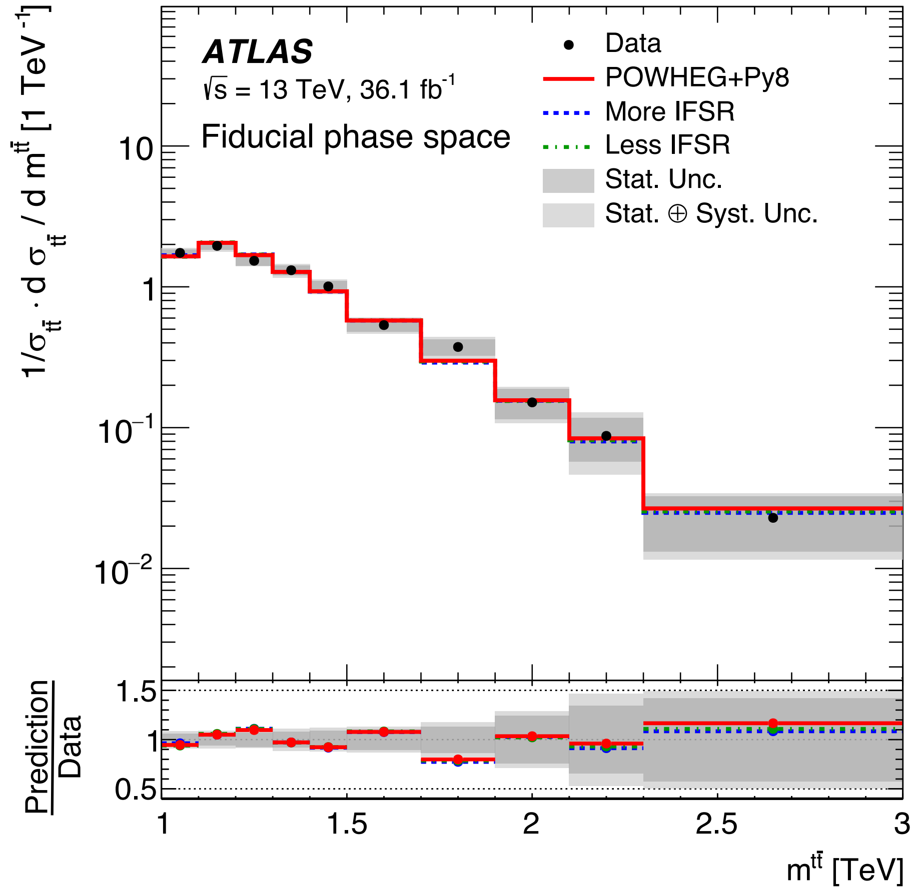




Unfolding Results: Cosθ* & m^{tt}



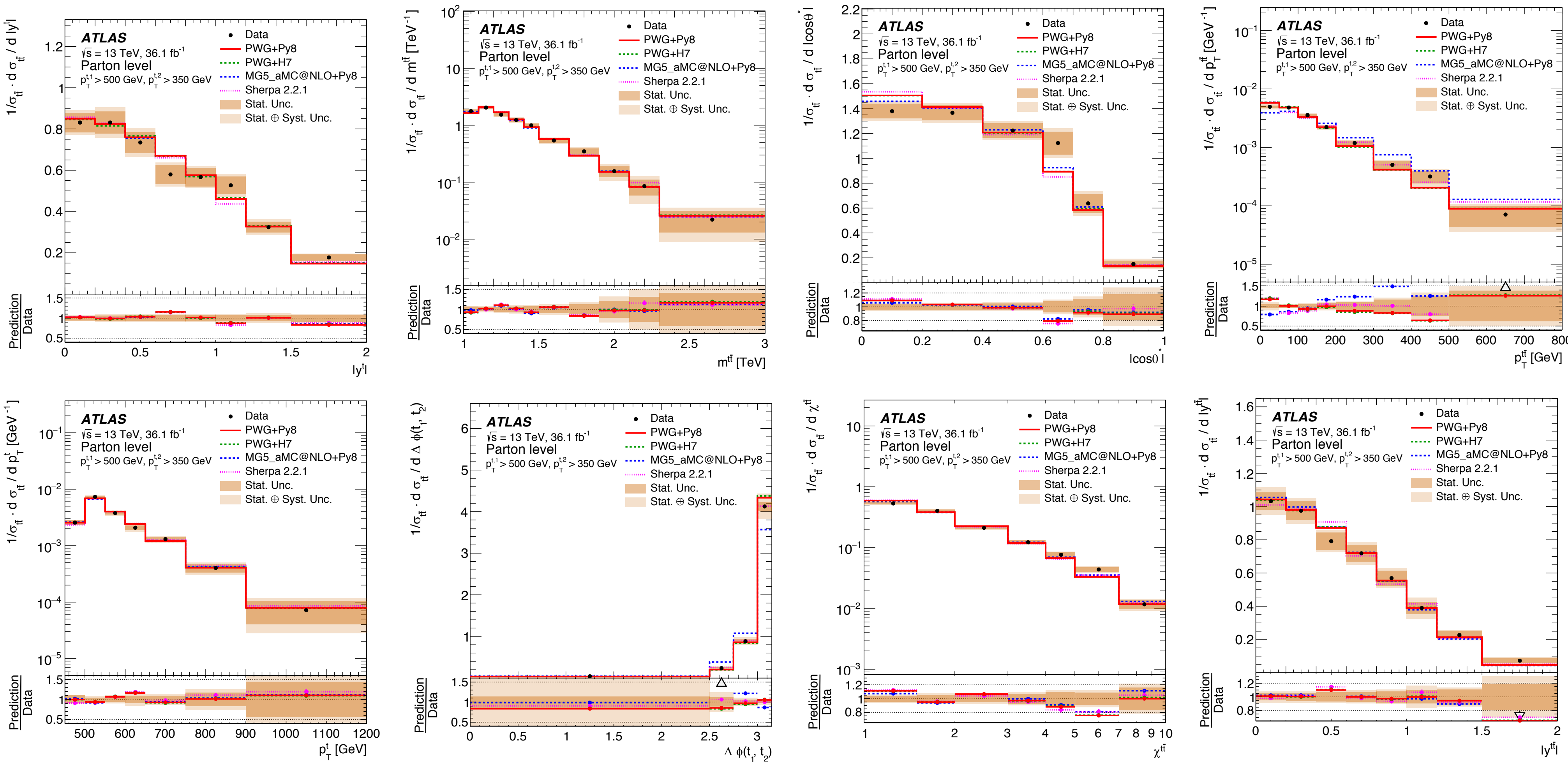




Unfold to Parton-Level

Repeat the procedure, unfolding to parton-level

- Allows for direct comparison with NLO and NNLO calculations

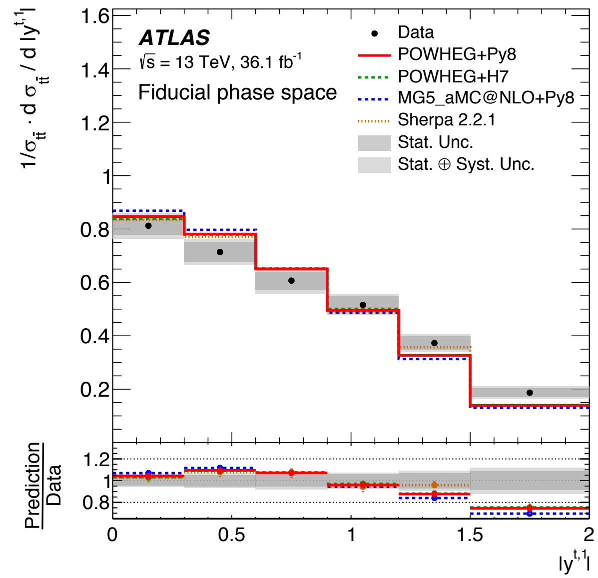


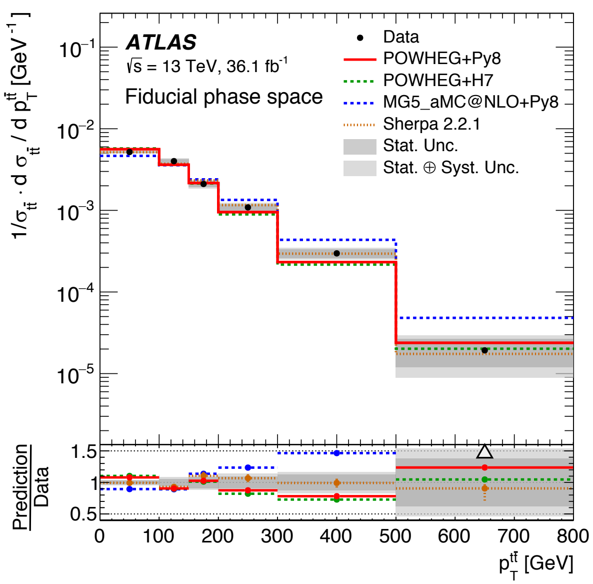
Compare with Theory Predictions

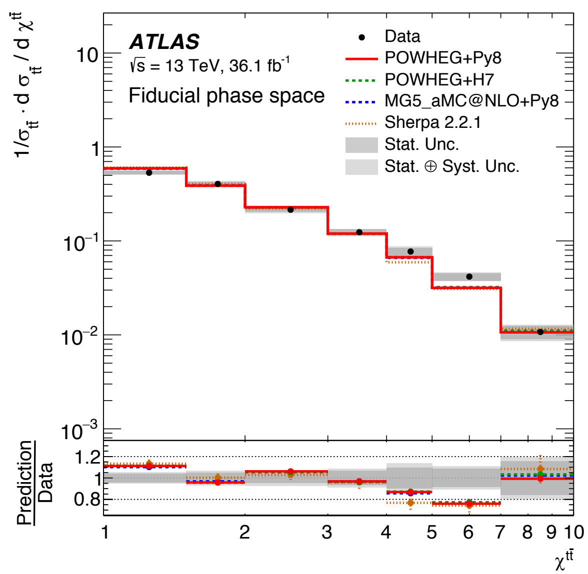
Generally good agreement

Several differences in differential cross sections

- Rapidity distribution of top jets broader
- p_T of top-pair system softer
- χ^{tt} distribution







$$\chi^{t\bar{t}} \equiv \exp(|\, y^{t,1} - y^{t,2}\,|)$$

Perform 1-D chisq Tests

Table 4: Comparison between the measured normalized particle-level fiducial phase-space differential cross-sections and the predictions from several SM event generators. For each variable and prediction, a χ^2 and a p-value are calculated using the covariance matrix described in the text, which includes all sources of uncertainty. The number of degrees of freedom (NDF) is equal to $N_b - 1$, where N_b is the number of bins in the distribution.

Observable	PWG+PY8		aMC@NLO +PY8		PWG+H7		PWG+PY8 (more IFSR)		PWG+PY8 (less IFSR)		SHERPA 2.2.1	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
$p_T^{t,1}$	7.7/7	0.36	8.2/7	0.32	8.0/7	0.33	9.1/7	0.24	8.7/7	0.27	9.3/7	0.23
$\|y^{t,1}\|$	7.5/5	0.18	12.2/5	0.03	6.8/5	0.24	8.8/5	0.12	8.1/5	0.15	4.0/5	0.55
$p_T^{t,2}$	8.6/6	0.20	2.6/6	0.86	9.9/6	0.13	12.2/6	0.06	5.0/6	0.54	5.0/6	0.55
$\|y^{t,2}\|$	3.7/5	0.59	4.6/5	0.46	3.1/5	0.68	3.5/5	0.63	3.2/5	0.67	2.9/5	0.72
$m^{t\bar{t}}$	4.5/9	0.88	4.7/9	0.86	4.0/9	0.91	5.3/9	0.81	5.2/9	0.82	10.0/9	0.35
$p_T^{t\bar{t}}$	7.8/5	0.17	20.9/5	<0.01	12.6/5	0.03	15.0/5	0.01	1.9/5	0.86	1.9/5	0.87
$\|y^{t\bar{t}}\|$	1.1/5	0.95	2.2/5	0.83	0.9/5	0.97	0.8/5	0.98	1.8/5	0.88	1.7/5	0.89
$\chi^{t\bar{t}}$	14.2/6	0.03	12.7/6	0.05	13.6/6	0.03	16.9/6	<0.01	10.1/6	0.12	18.5/6	<0.01
$y_B^{t\bar{t}}$	2.5/6	0.87	3.3/6	0.77	2.2/6	0.90	2.6/6	0.86	2.8/6	0.84	3.0/6	0.81
$\|p_{out}^{t\bar{t}}\|$	1.9/6	0.93	53.1/6	<0.01	3.1/6	0.80	4.2/6	0.64	4.8/6	0.57	5.9/6	0.44
$\Delta\phi^{t\bar{t}}$	0.9/3	0.84	16.3/3	<0.01	2.0/3	0.58	3.0/3	0.40	0.6/3	0.89	3.4/3	0.33
$H_T^{t\bar{t}}$	4.8/6	0.57	5.2/6	0.52	4.5/6	0.61	5.0/6	0.54	5.0/6	0.55	3.1/6	0.80
$\cos\theta^\star$	9.9/5	0.08	10.5/5	0.06	9.3/5	0.10	12.8/5	0.03	6.5/5	0.26	18.7/5	<0.01

More Observations

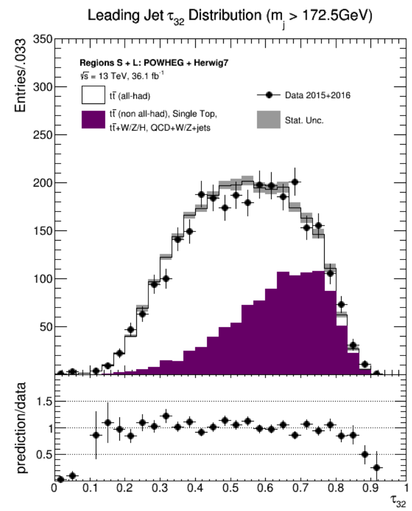
Efficiency sensitive to hadronization – 10-15%

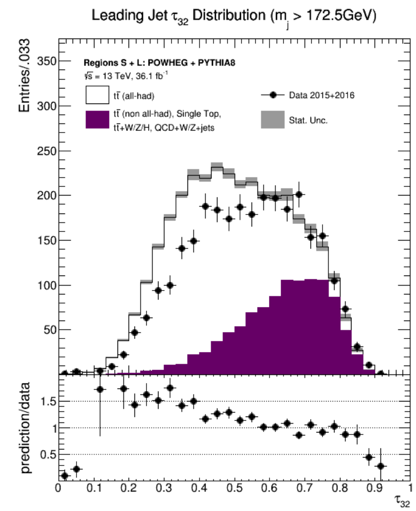
- Discovered that τ_{32} most sensitive
- Directly affects the efficiency of top-tagging

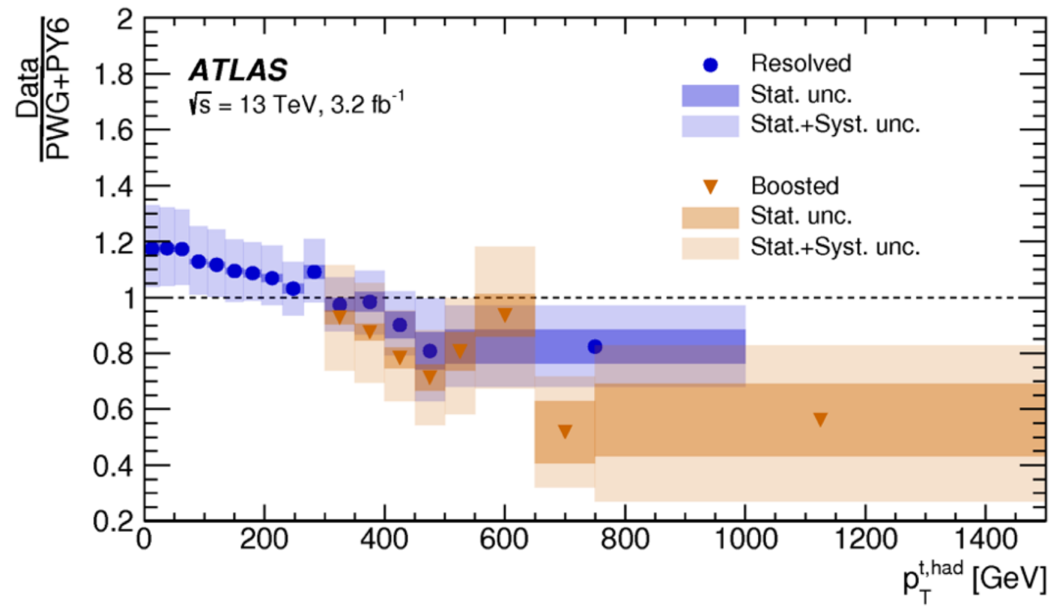
Models predict harder top jet spectrum

- Confirms measurements at lower p_T

JHEP 11 (2017) 191

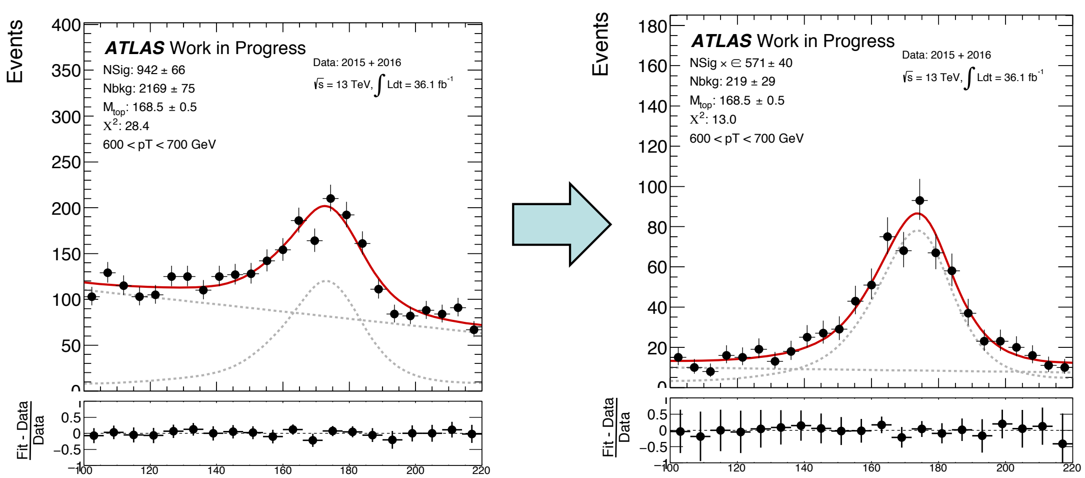






Moving to Greater Precision

- **Increase data sample**
  - Working with 2017 data → x2 increase
  - Started collisions last week → another x2?
  - Measure doubly-differential cross sections



- **Systematic uncertainties**
  - B-tagging – using top mass as a tool to constrain uncertainty
  - Parton-shower & hadronization
    - Exploring the difference in models
    - Improve the models and/or make tagging less sensitive to differences?
  - ME modelling
    - Ongoing effort with MC developers to understand POWHEG and aMC@NLO differences

Summary and Conclusions

Completed first differential cross-section measurement using boosted top jets

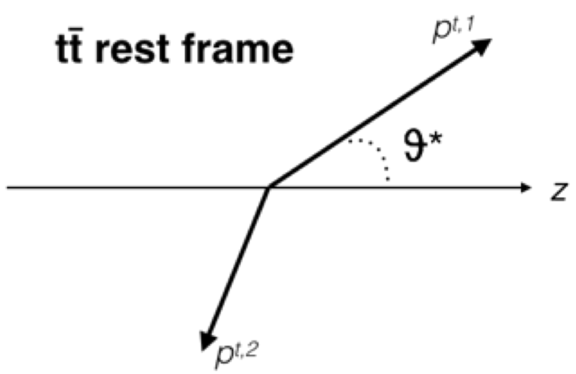
- Exploring mass scales > 2 TeV
- Revealed generally good agreement with QCD predictions
- Softer production than expected
- More top pairs at large angles

Discriminating between theory predictions

- Focuses activity on understanding physics differences

Expect to make improved measurements over next 2 years

Submitted to PRD: arXiv:1801.02052



BackUp Stuff

Friday September 06, 2013

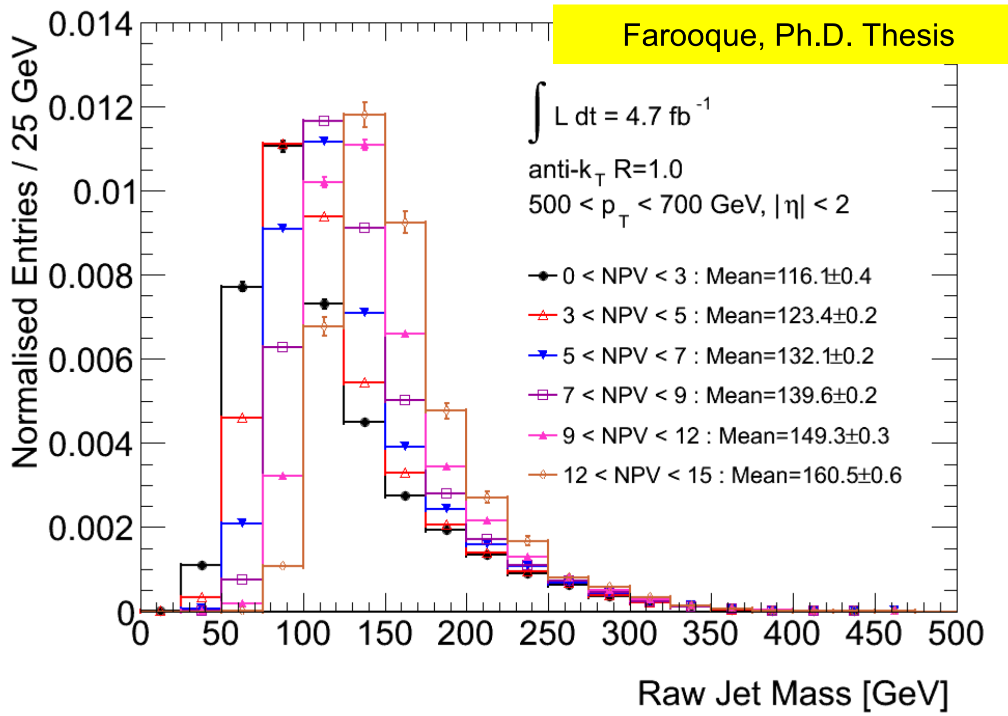
Addressing Pile-Up

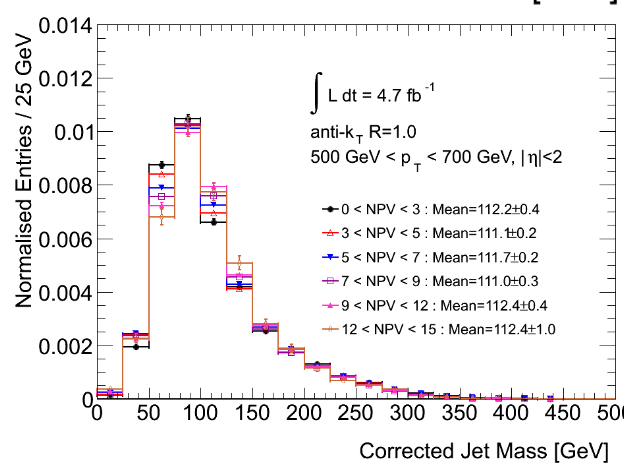
Jets are extended objects

- Contributions from additional interactions have significant effect on observed properties

Various strategies to address

- Correct with average calibration
  - Only used at Tevatron, and never on jet substructure
- Correct event-by-event
  - Works OK but cumbersome
- Can "cut-out" pile-up contributions
  - This is method of choice
  - Requires careful calibration



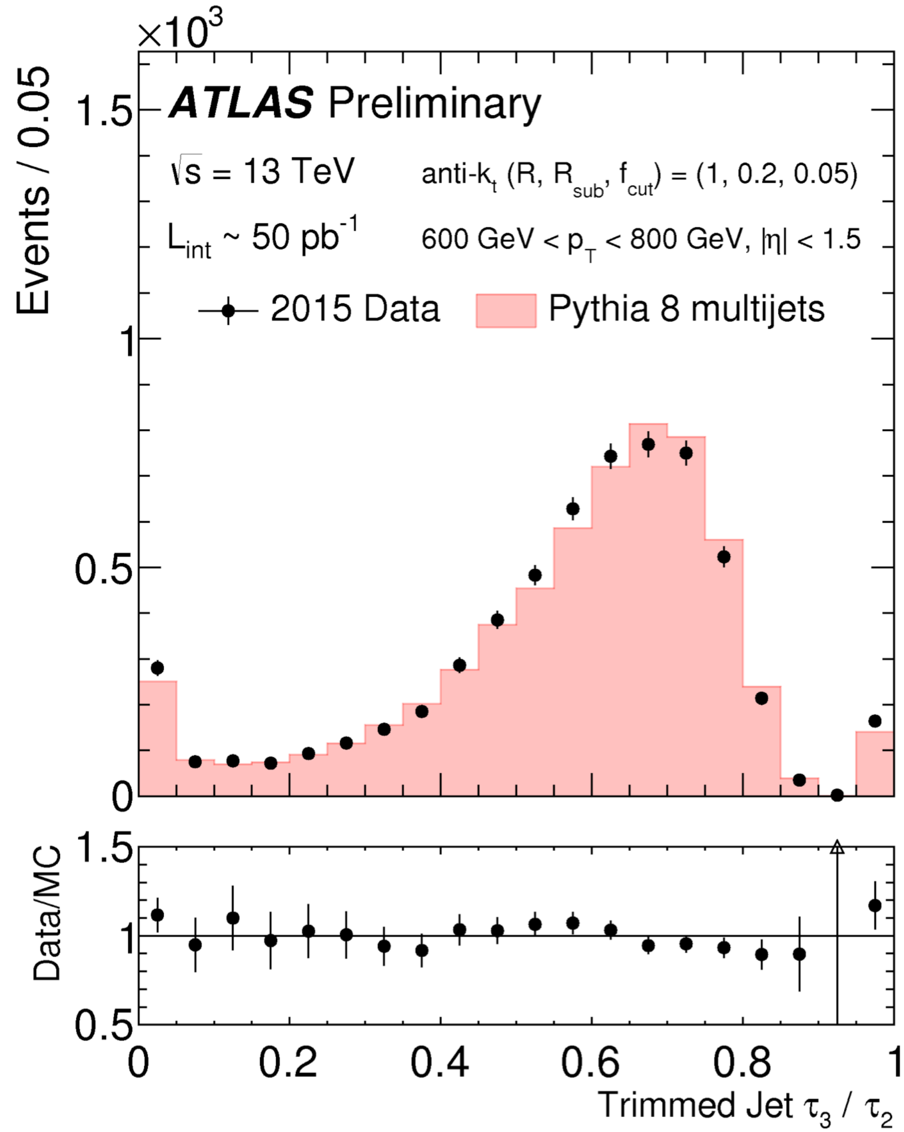


Tau$_{32}$ Modelling

$$\tau_3 = \frac{1}{d_0}\sum_{k=1}^{3} p_{Tk} \times \min(\delta_{1k},\delta_{2k},\delta_{3k})$$

$$\tau_2 = \frac{1}{d_0}\sum_{k=1}^{2} p_{Tk} \times \min(\delta_{1k},\delta_{2k})$$

$$d_0 = \sum_{k=1}^{3} p_{Tk} \times R$$



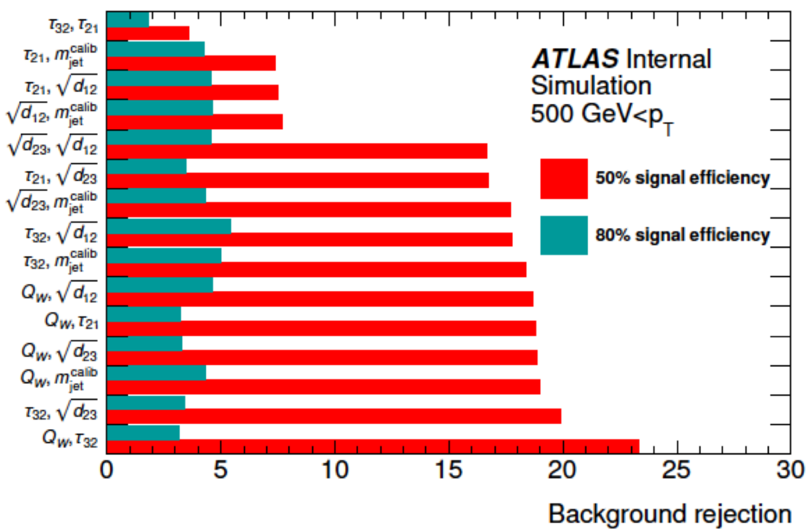
Thaler & van Tilburg, JHEP 03 (2011) 015; JHEP 02 (2012) 093

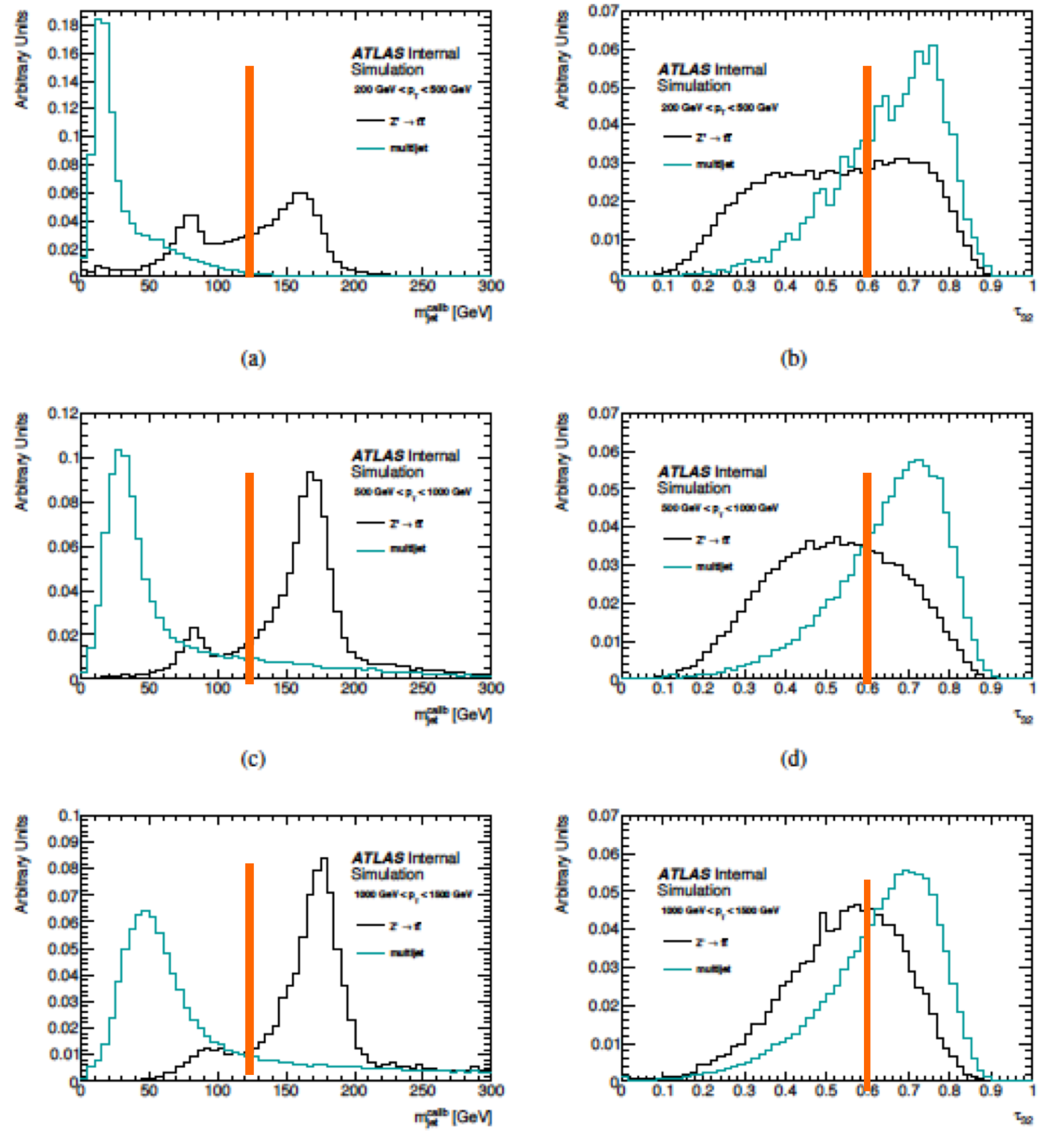
Current Choice of Algorithms

Looked at algorithm with 2 variables:

Optimitized for jets with $p_T > 500$ GeV,

- $M_{jet} > 125$ GeV
- $\tau_{32} > 0.58$





Aside: Trimming Cuts Out QCD Too

Trimming removes part of the QCD jet as well

- Current parameters remove 100 GeV subjet for a 1 TeV object!
- Requires good fragmentation models in order to calibrate correctly
- Competing schemes for this
  - Mass-drop
  - C/A clustering
- Optimizing f_{cut} and R_{sub} remain open questions