Jet substructure and ttbar resonance searches with boosted tops

Elena Oliver, Miguel Villaplana, Marcel Vos, Santiago Gonzalez de la Hoz, José Salt, Eduardo Ros, Victoria Sánchez

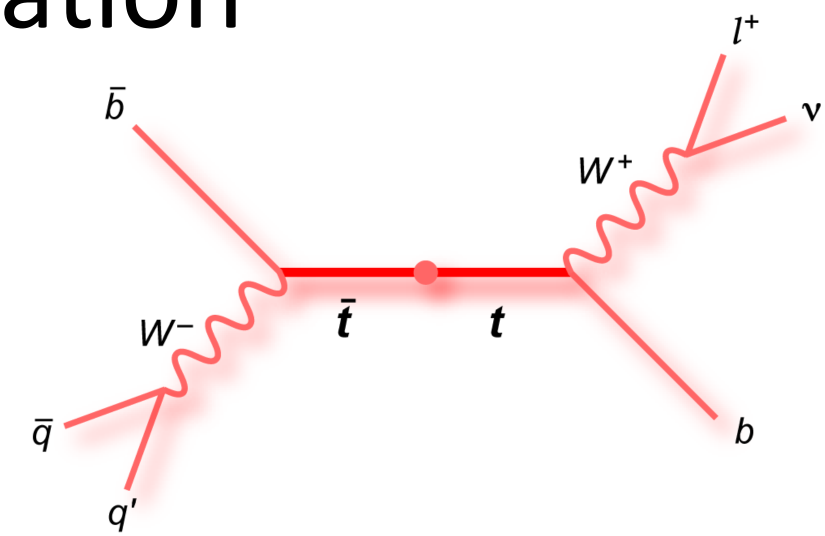
I PCI2011 Workshop
Rabat 27 June 2012

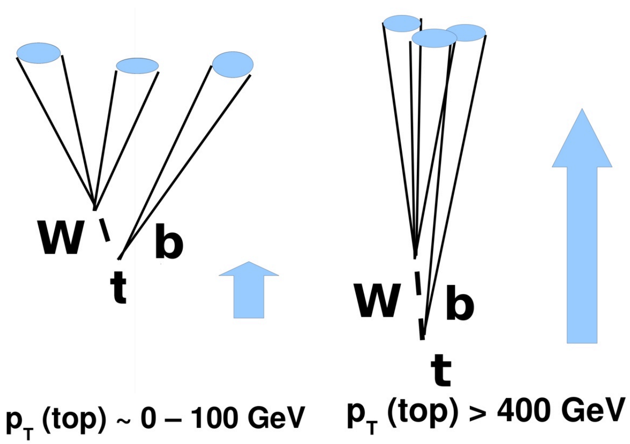
Overview

- Motivation
- Jet Algorithms
- Jet techniques for boosted tops
- Commissioning jet substructure:
  - Unfolding
  - Results
- ttbar resonance searches with boosted tops
  - Boosted top quark reconstruction
  - Background
  - ttbar distribution
  - Searches
- Conclusions

Motivation

- The top quark plays a special role in many BSM proposals. For example:
  - strong coupling,
  - produces isolated leptons and can be easily distinguished from the anti-top.
- At LHC energy scales, particles such top quarks are produced abundantly with significant Lorentz **boosts.**





When these tops decay hadronically, the products are collimated in a small area of the detector, if large boosts, resulting hadrons can be clustered into a **single jet.**

Substructure can extract the single jets of interest from the overall jet background.

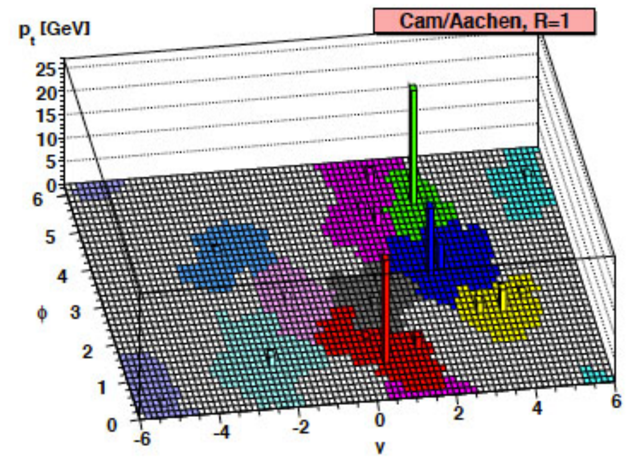
Jet Algorithms

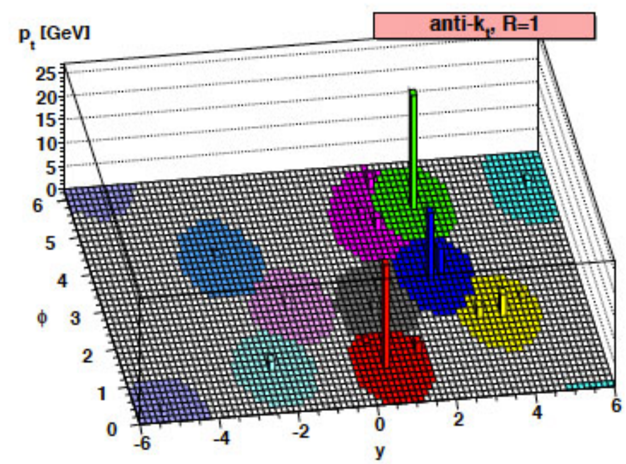
- **Jet algorithms** combine the clusters into jets, according to their p_T and relative separation
- The resulting jets are formed with a pre-defined radius, R, in η-ϕ space.

- Inclusive jet algorithms:
- $d_{ij}=\min(k_{ti}^{2p}, k_{tj}^{2p})\,\Delta R_{ij}^2/R^2$
- $d_{iB}=k_{ti}^{2p}$

- 1º Work out all the d_{ij} and d_{iB}
- 2º Find the minimum of the d_{ij} and d_{iB}
- 3º If minimum d_{ij} -> recombine I and j, go to 1º
- 4º If minimum d_{iB} -> i considered jet, go to 1º without i
- 5º Stop when no particles remain

- For Kt Algorithm -> p=1
- Cambridge-Aachen (p=0) :Closest constituents clustered first.
- Anti kt (p=-1) : Hardest constituents combined first. Gives circular jets resilient to soft radiation.

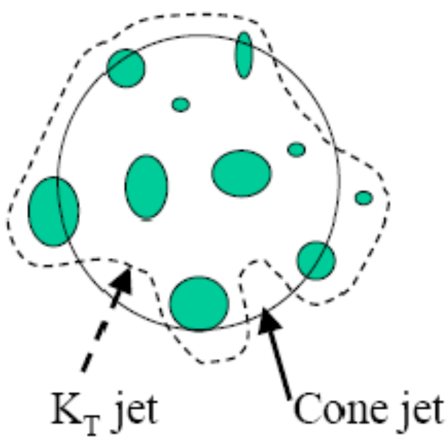




Ref: arXiv:0906.1833v2, arXiv:0802.1189

Jet techniques for boosted tops

- ATLAS jet finding default is anti-k_T (R=0.4 or 0.6)
- For a parent particle with m and p_T, merging starts showing at $R > 2m/p_T$
- Use jet mass and jet substructure to resolve merging
- Rerun jet algorithms on jet components to reveal jet substructure (k_T or Cone)

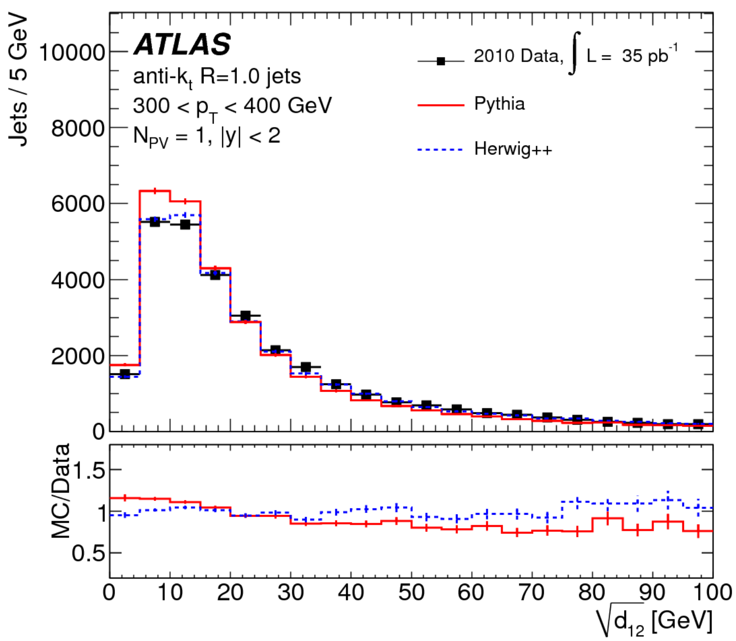


- Jet Mass: sum over mass-less four-vectors of jet components.
- Kt splitting scales $\sqrt{d_{ij}}$: value before re-clustering the constituents of the jet with the kt recombination algorithm
  - $\sqrt{d_{12}} = \min(p_{Tj_1}, p_{Tj_2}) \times \delta R_{j_1,j_2}$ in the last clustering step
  - $\sqrt{d_{23}}$ in the next-to-last clustering step

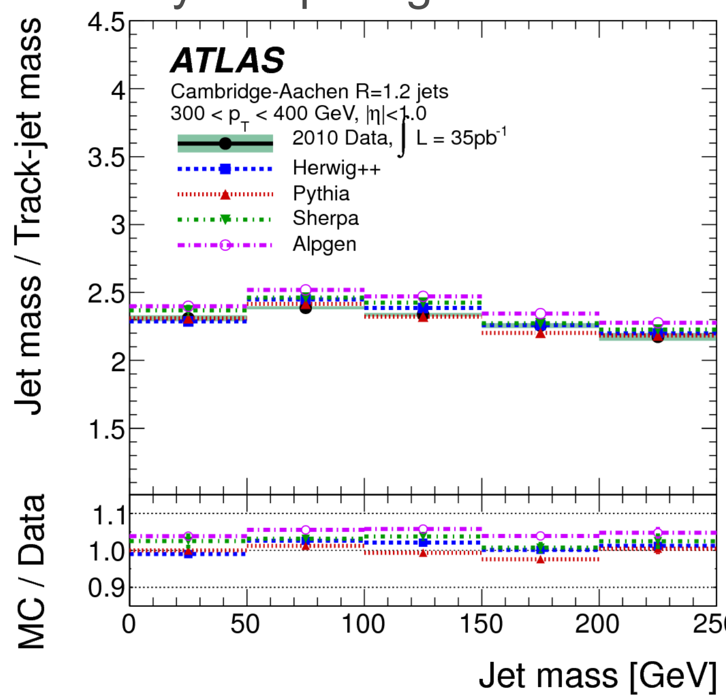
Comissioning jet substructure

- The energy and mass calibration and scale uncertainties for the anti-k_t jets with R=1 are presented in **arXiv:1203.4606 [hep-ex]**
- Data from ATLAS 2010 pp collisions at $\sqrt{s}$=7 TeV, Luminosity= 35 pb^{-1}
- Use locally calibrated topological clusters so that jet-level corrections are small.
- Determine jet-level corrections on MC to correct reconstructed energy or mass to the scale of matched particle jets

Check the detector-level distribution



Determine scale uncertainty for E, m and $\sqrt{d_{12}}$ in situ by comparing track and calo jets



Comissioning jet substructure: Unfolding

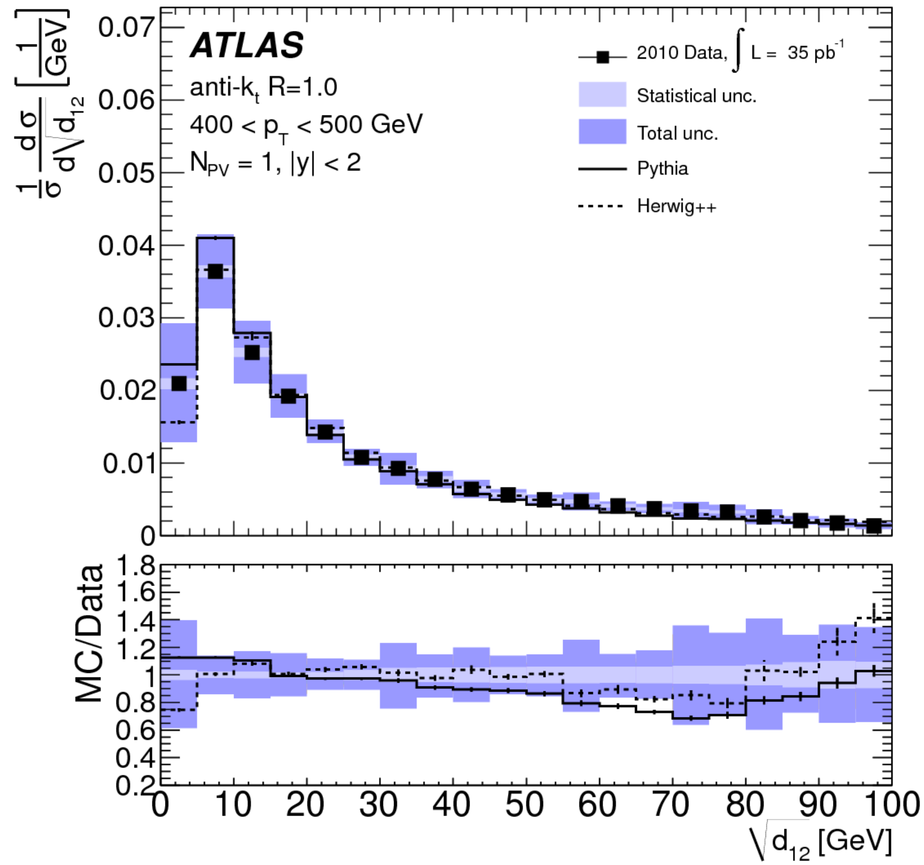
- To compare measurements and theoretical predictions -> corrections for detector resolution
- Unfolding type: Interactive Dynamically Stabilised (IDS)
- With Monte Carlo, comparing truth and reconstruction jets and obtaining a transfer matrix
- The data are then scaled by the reconstructed matching efficiency, multiplied by the transfer matrix and divided by the truth matching efficiency.
- Each p_T bin is unfolded independently.

Comissioning jet substructure: Results

arXiv:1203.4606 [hep-ex]

Distribution at "particle-level", after unfolding show reasonable agreement between data and most MC within not-too-large systematic uncertainties:

- Parton Shower model is adequate
- Detector response is under control
- Underlying event OK



Pile-up has been shown to have a big impact on some substructure observables (most notoriously, jet mass)

We can mitigate the impact on analysis by grooming or smart choice of observables, by correcting using smart techniques, and by modeling pile-up correctly in MC

ttbar resonance searches with boosted tops

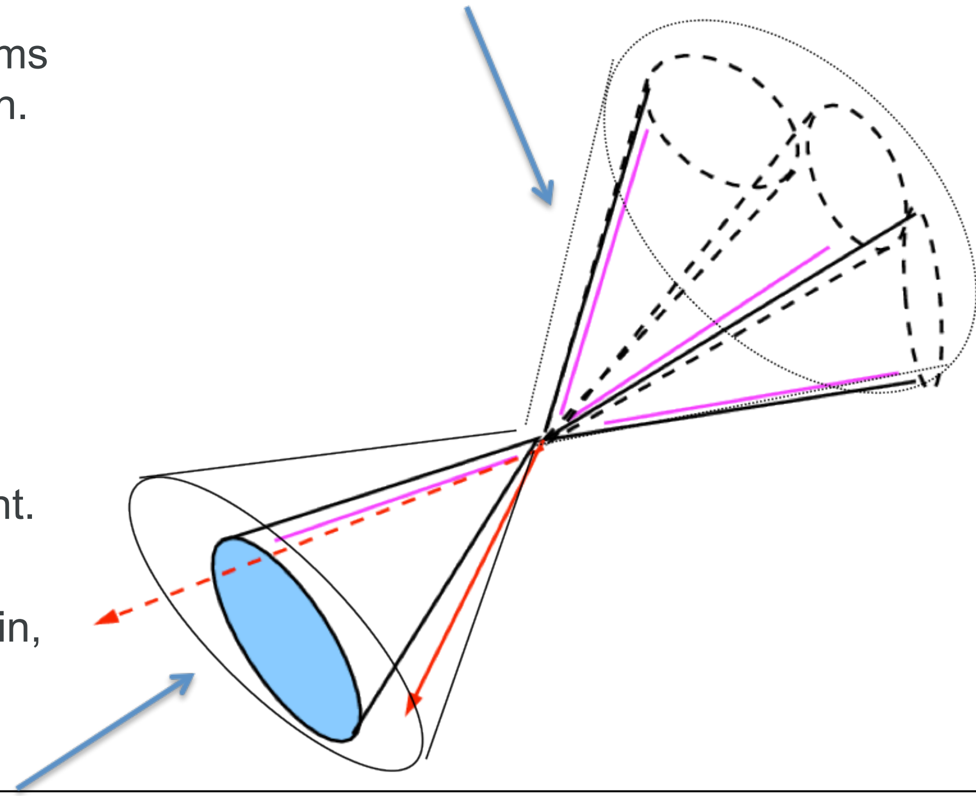
- Publication: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TOPQ-2011-23/TOPQ-2011-23.pdf
- Data from ATLAS 2011 pp collisions at $\sqrt{s}$=7 TeV, Luminosity= 2.05 fb^{-1}

Boosted top quark reconstruction

t → bW → bjj
reconstruct a single "fat" jet (R=1-1.5)
measurable substructure (jet mass, splitting scales, ...)

Top reconstruction algorithms well-established at Tevatron. Similar for ATLAS top physics.

- Find isolated lepton.
- Find 4 jets.
- Reconstruct neutrino using lepton, Emiss
- T and W-mass constraint.
- (Optional) Tag b-jets.
- Piece it all together again, selecting the right combination.



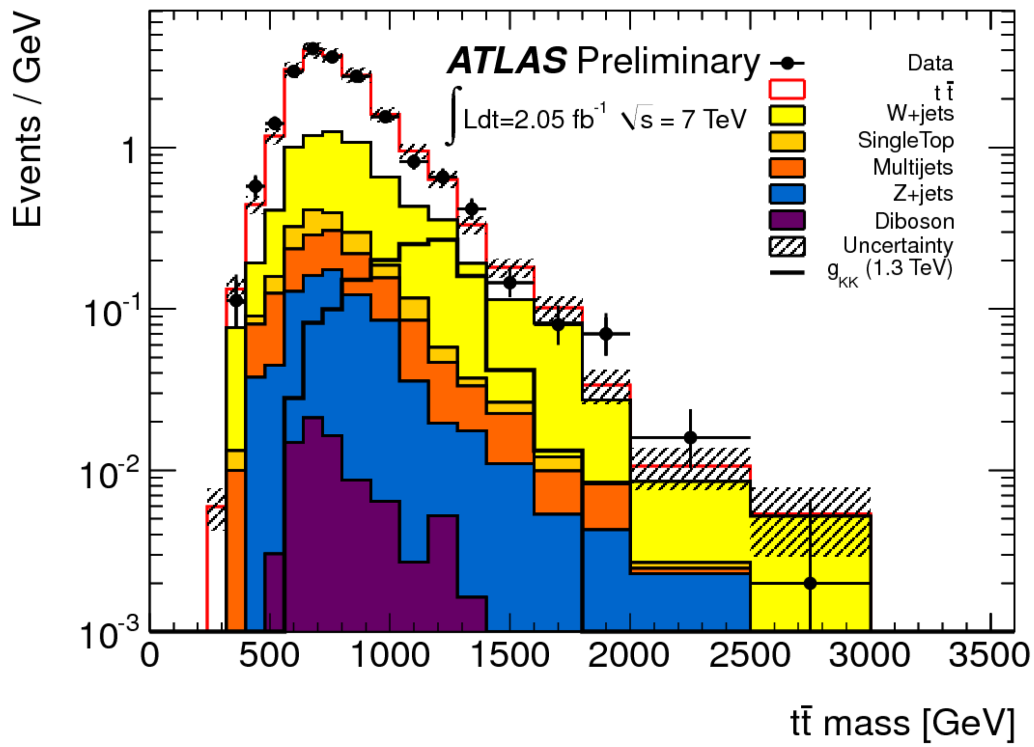
t → bW → blν
Reconstruct with antiKT R=0.4 (small area is expected)

Background

Type	e+jets	μ+jets	Sum
$t\bar{t}$	510 ± 40	620 ± 50	1130 ± 90
W+jets	202 ± 34	300 ± 50	500 ± 80
Multijets	45 ± 23	30 ± 15	75 ± 38
Z+jets	41 ± 20	34 ± 16	75 ± 36
Single top	21 ± 2	27 ± 3	48 ± 5
Dibosons	3.4 ± 0.2	4.5 ± 0.2	7.9 ± 0.4
Total	830 ± 60	1010 ± 70	1840 ± 130
Data	803	1034	1837

Selected data events and expected background yields after the full selection. The statistical uncertainty on the observed number of events and the uncertainties on the normalization of the expected background yield are listed

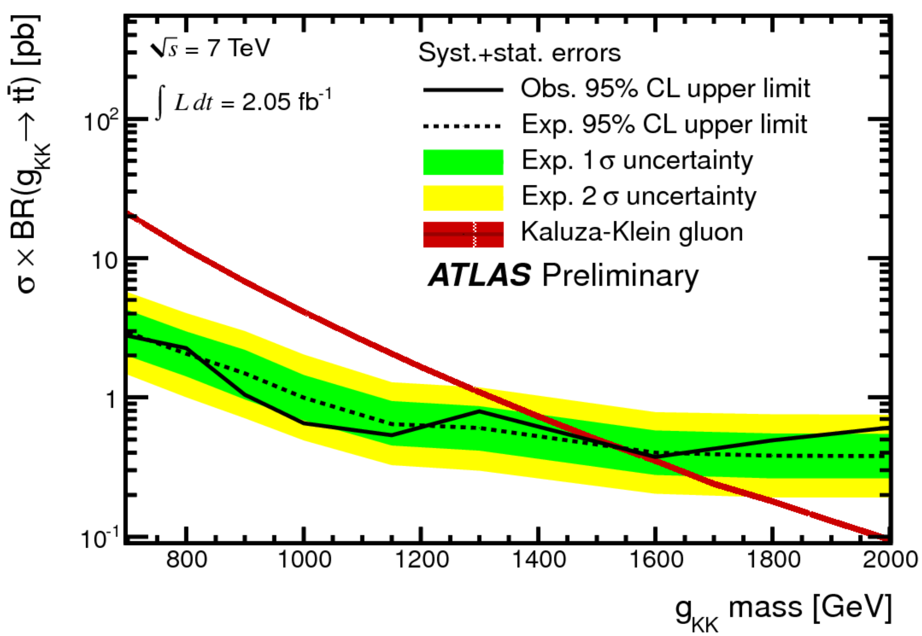
ttbar mass distribution

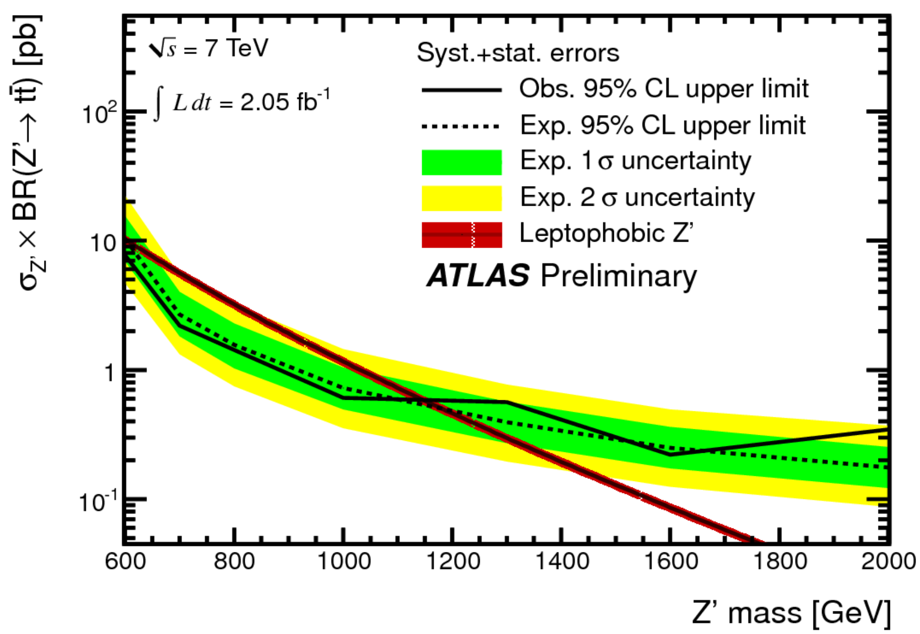


Reconstructed invariant mass distribution of the ttbar pair after the signal selection. The shaded band indicates the uncertainty in the normalization of the Standard Model prediction, but does not include the shape uncertainty or the impact of uncertainties on reconstructed objects. The variable bin size is chosen to match the mass resolution for a resonant signal.

Searches

Rule out (at 95% C.L.) the existence of a narrow (leptophobic) Z' (in topcolor models) or a heavy broad KK excited state of the gluon.





Conclusions

- The LHC in the boosted regime
- Boosted top quarks behave as expected
- Algorithms designed specifically for the topology of boosted top quarks are expected to “boost” our physics objective
- ATLAS has a fully commissioned top-tagger for hadronic decays of boosted top quark and is preparing more sophisticated tools
- ATLAS limits from searches using boosted techniques are the best showed so far (better than ATLAS resolved and than CMS results)

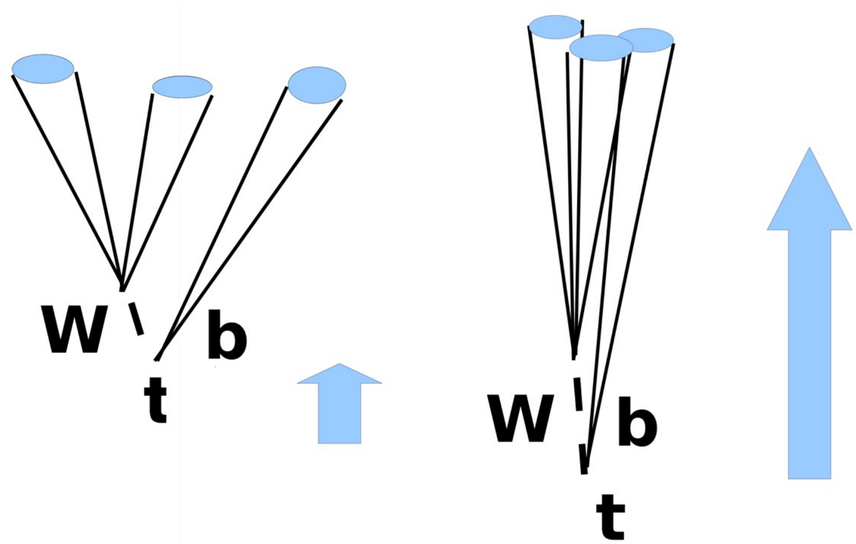
References

- Notes:
  - https://cdsweb.cern.ch/record/1421216
  - https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TOPQ-2011-23/TOPQ-2011-23.pdf
- Posters:
  - https://cdsweb.cern.ch/record/1386662/files/ATL-PHYS-SLIDE-2011-604.pdf
  - https://cdsweb.cern.ch/record/1374657/files/ATL-SOFT-SLIDE-2011-449.pdf

Backups

Boosted Object Definition

- **Let's define "boosted object" by comparing the standard approach** (reconstruct components and combine) **to Mike Seymour's alternative** (find composite object and decompose).
- **Rules of thumb** for maximum jet radius parameter for 2-body decay:
  - $R < 2m/p_T$ (always resolve two jets)
  - $R > 3m/p_T$ (capture full decay in a single jet 75% of cases)

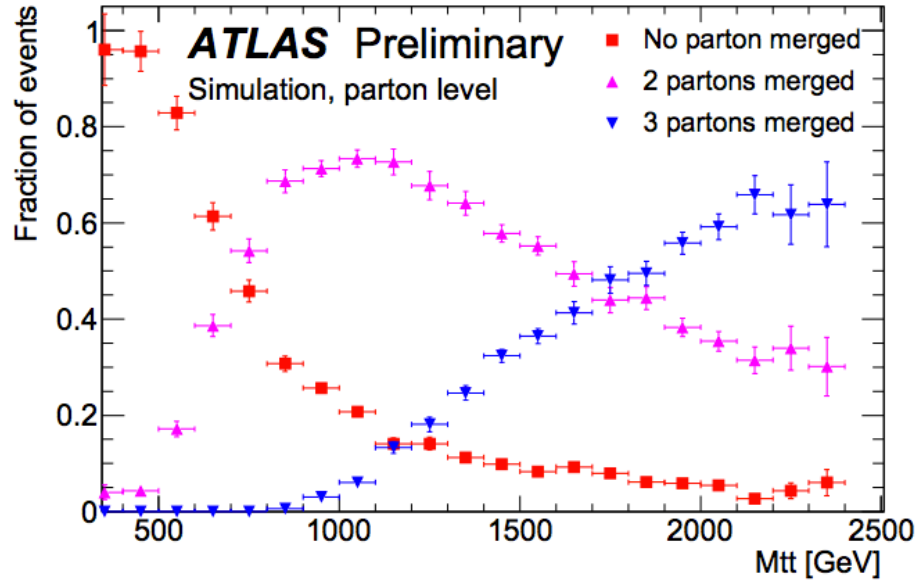


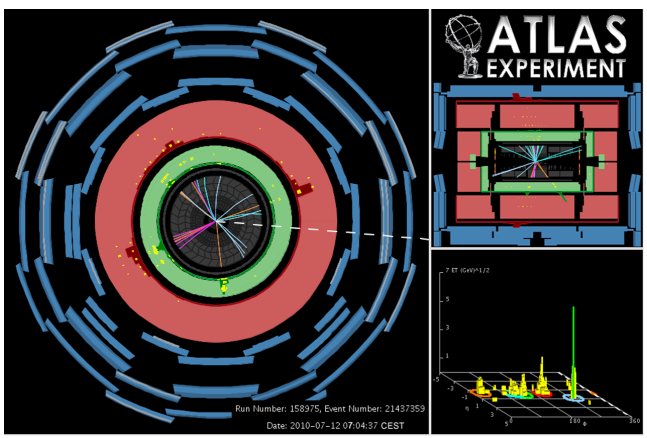
p_T (top) ~ 0 – 100 GeV

p_T (top) > 400 GeV

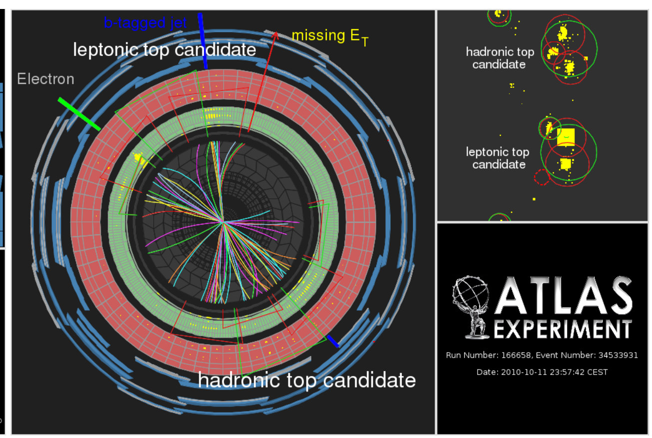
Top Reconstruction: Three regimes

- W boson at rest
  - use resolved approach
- $p_T \sim 240$ GeV
  - coexisting algorithms, can resolve with R=0.4, or contain in R=1
- $p_T \sim 400$ GeV
  - boosted regime cannot always resolve with R=0.4

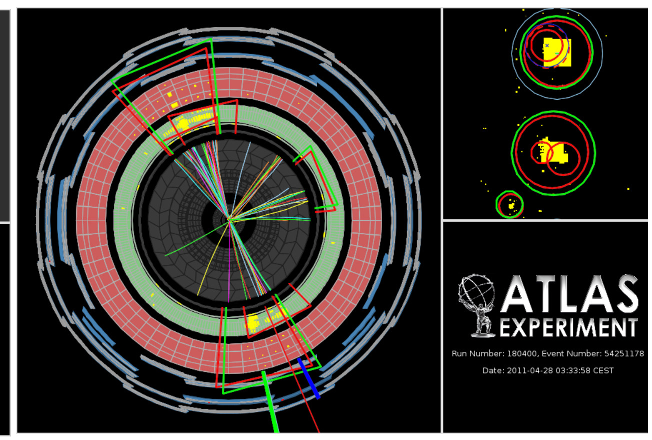




Early "l+jets" candidate
ATLAS-CONF-2010-063



First boosted top quark
ATLAS-CONF-2011-073



$m_{tt} > 1$ TeV
ATLAS-CONF-2011-083

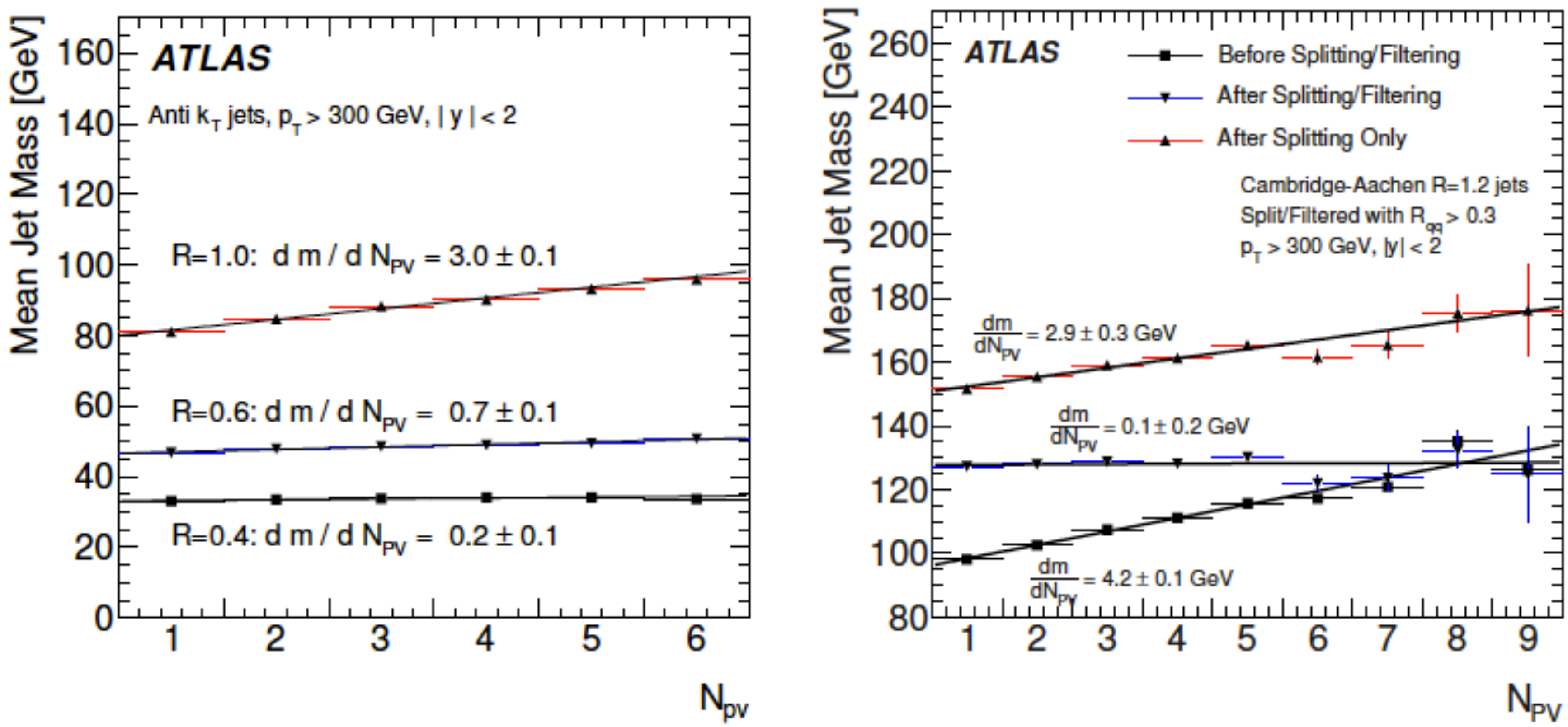
JS:Monte Carlo Samples

- NO PILE-UP
- Several MC generators for inclusive jets (LO,pQCD):
  - PYTHIA 6.423
  - HERWIG++ 2.4
- MC generators for cross-checks:
  - ALPGEN 2.13<-HERWING++
  - SHERPA 1.2.3
- Parton-shower LLogO
  - PYTHIA by PT ordered
  - HERWIG++ by angular ordered
- Fragmentation into particles
  - PYTHIA by following the string
  - HERWIG++ following cluster model
- Tunes:
  - PYTHIA: AMBT1, Perugia 2010, MRST2007 LO*

JS: Calibrations and Scale uncertainties

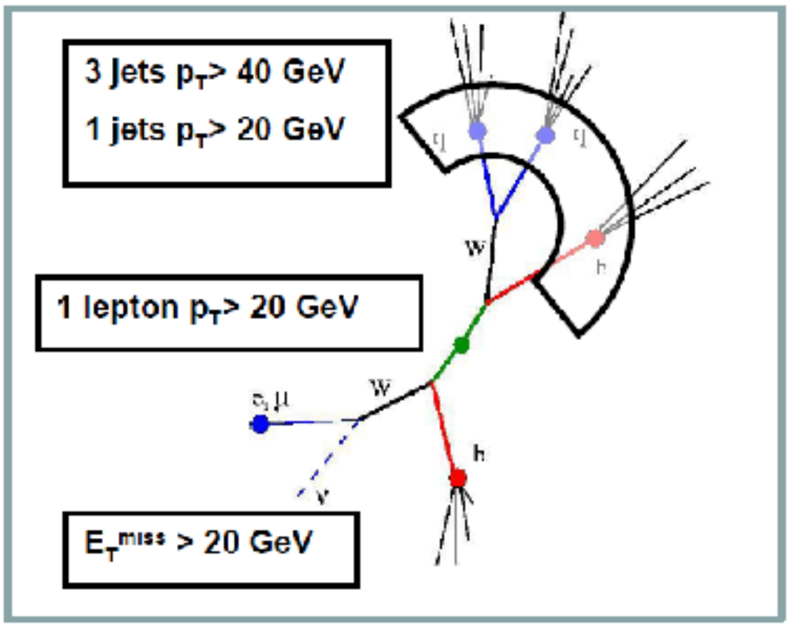
- Locally Calibrated TopoClusters used (different from ATLAS) -> Cluster's energies are corrected with calibration constants for smaller jets.
- Because larger R-parameters, ATLAS jet energy scale uncertainty for anti-kt R = 0.4 and 0.6 jets cannot be applied.
- Jet constructed from tracks are used for systematic studies (track-jets in inner-detector only from the selected pp collision of interest: pT > 500 MeV and |z0| < 5 mm, z0 is the z-coordinate of the track at closest approach to the z-axis.). These trackjets are constructed using the same algorithms as calorimeter jets.
- For p_T and each substructure variable, obtaining uncertainties in scale and resolution. Because Inner Detector and calorimeter have uncorrelated systematic effects -> some separation of physics and detector effects.
- Scale uncertainties: Tracks-jets matched to calorimeter-jets if $\delta R<0.3$
- $r^X = X_{calorimeterjet}/X_{trackjet}$ (X each variable)
- To quantify the level of agreement: $\rho^X = r^X_{data}/r^X_{MC}$

JS: Pile-up Dependency

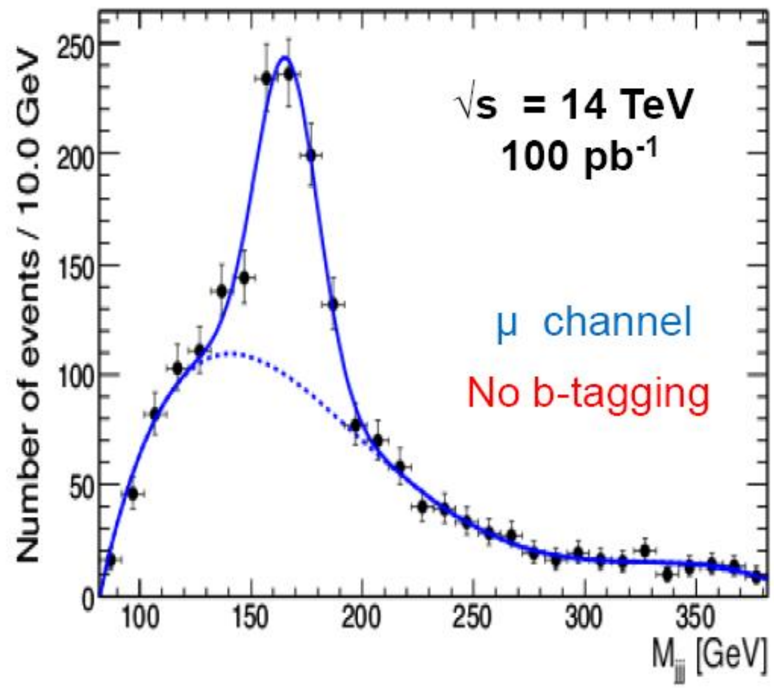


Top reconstruction basics

- Top reconstruction algorithms are well-established at the Tevatron. The same "resolved" approach forms the back-bone of the ATLAS preparation for top physics.
- Find isolated lepton.
- Find 4 jets.
- Reconstruct neutrino using lepton, Emiss
- T and W-mass constraint.
- (Optional) Tag b-jets.
- Piece it all together again, selecting the right combination.



tt -> Wb Wb -> ℓνb qqb



three jet mass

Are boosted top quarks particularly relevant for the LHC program?

- Hoping to directly observe a new source:
  - something at the TeraScale (e.g. Z') decays to something at the electro-weak scale (t,W/Z,H,…) that decays to N jets

  $X \rightarrow tt$, $b' \rightarrow tW$, $W' \rightarrow tb...$, $g \rightarrow tt\, \chi_0$,

- Even for non-resonant production, one may look for a region with less background and/or ambiguities

  ttH → Plehn, Salam, Spannowsky '09

- Many handles on study of top quark production:
  - same-sign tops,
  - charge asymmetry
  - polarization
  - spin correlations

Bt: Object Selection

- For the jets, the object selection cuts are
  - **akt4:** |eta|<2.5, pT>30 GeV, E>0
  - **akt10:** |eta|<2.0, pT>150 GeV, E>0 (following substructure note recommendations).

- Cut flow for event selection
  - **Pre-selection: up to and including C9 of the top cross section cut flow.**
  - **Cut 0:** require ELECTRON trigger for the e+jets and MUON for the mu+jets channels
  - **Cut 1:** Non-collision background rejection: Require that the first primary vertex in the VxPrimaryCandidate has Ntracks>4 and is of type PriVtx or type PileUp
  - **Cut 2:** At least 1 electron ET>25 GeV (e+jets) or 1 muon PT>20 GeV (mu+jets)
  - **Cut 3:** Exactly 1 electron ET>25 GeV (e+jets) or exactly 1 muon PT>20 GeV (mu+jets)
  - **Cut 4:** Exactly 0 muons PT>20 GeV (e+jets) or exactly 0 electrons ET>25 GeV (mu+jets)
  - **Cut 5:** the lepton matches the trigger
  - **Cut 6:** remove events tagged as e-mu overlap
  - **Cut 7:** Jet Cleaning (for data ONLY): no bad jets with Pt > 20 GeV
  - **Cut 8:** Transverse missing energy (MET) cut -- e+jets: MET>35 GeV, mu+jets: MET>20GeV
  - **Cut 9:** W tranverse mass (MTW) cut -- e+jets: MTW>25 GeV, mu+jets: MET+MTW> 60 GeV

Bt: Object Selection

- **Cut 10':** Remove the akt10 jet closest to the electron if it also fulfils dR(ele, akt10 jet)<0.5. Require at least one akt10 jet with pt > 150 GeV.

- **Cut 11':** Require at least one akt4 jet with 0.4 <dR(lep,jet) <1.5.

- Choose the candidate for the b-jet from the leptonic top decay from the akt4 jets that fulfil the above cut (0.4 <dR(lep,jet) <1.5). Pick the **jet closest to the lepton**. This is **the leptonic top** jet.

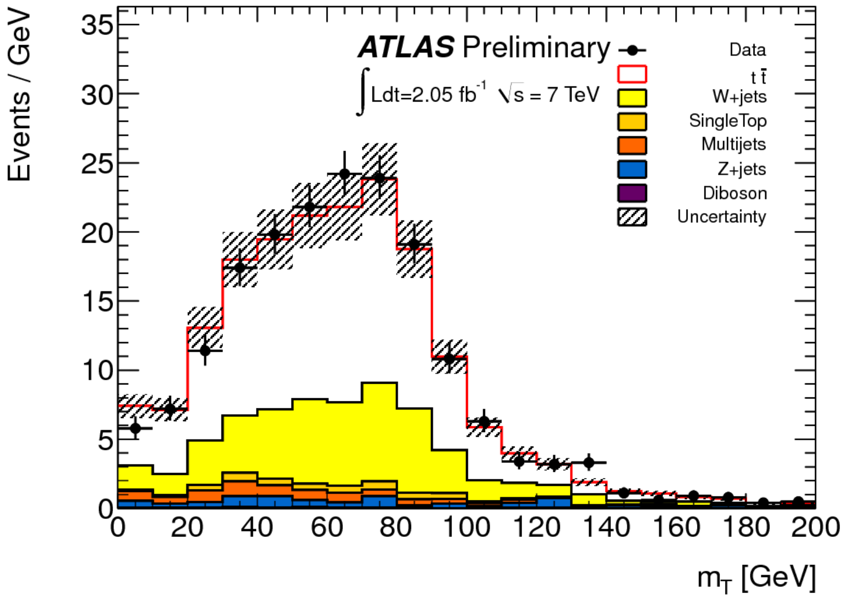
- **Cut 12':** Remove akt10 jets overlapping with the leptonic top jet, dR(lep top jet, akt10)<1.5. Remove all akt4 jets except the leptonic top jet. Require at least one akt10 jet with:
  - akt10_pt > 250 GeV
  - akt10_m > 100 GeV
  - akt10_split12 ($\sqrt{d_{12}}$) > 40 GeV

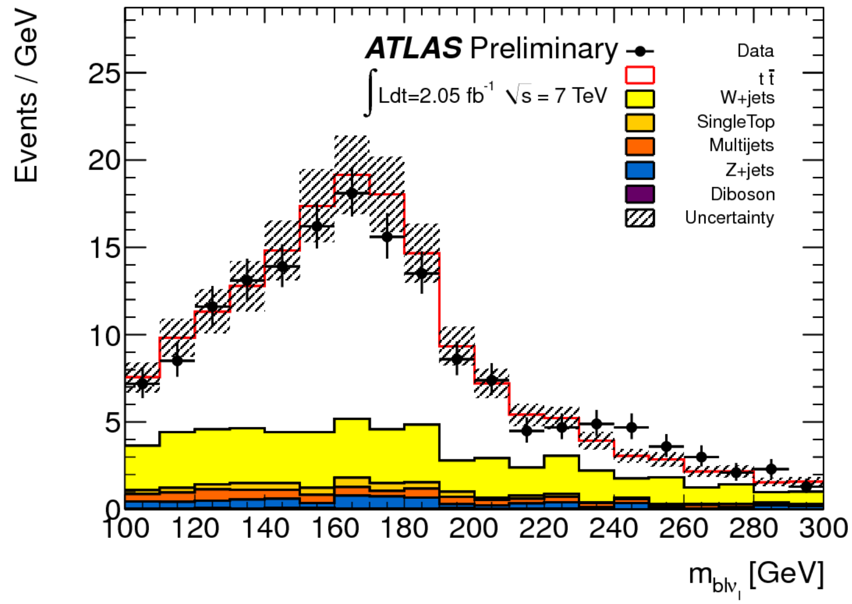
- As **hadronic top jet** choose the **akt10 jet surviving C12' that has the highest pt**.

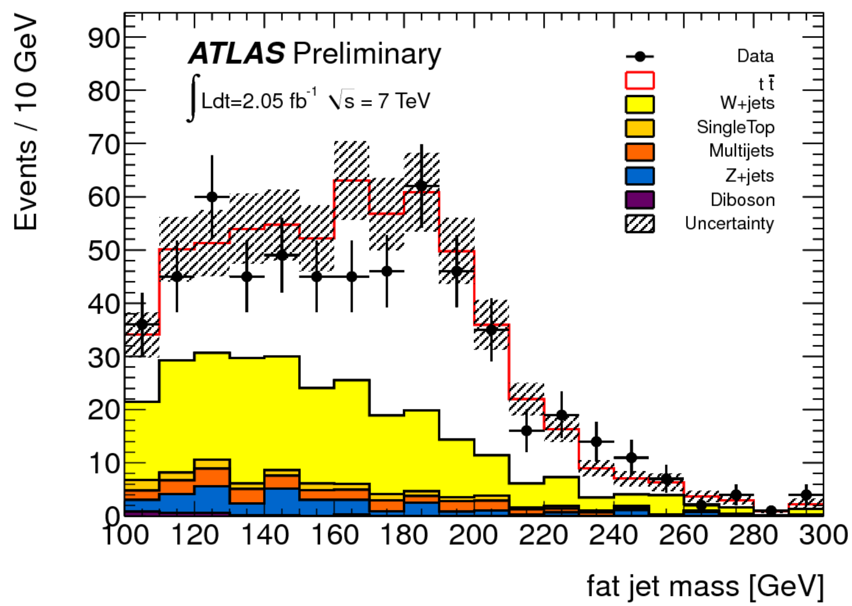
- **Cut 14:** (from standard cut flow - implement on DATA only): LAr error flag cut:
  larError ! = 0

- As a jet cleaning cut, we require that we have no bad akt4 jets in the event (this is to clean also the akt10 sample). This is already done in the standard top recommendations.

Bt: Some control plots







Bt: Systematics

Systematic effect	Impact on yield [%] background	Impact on yield [%] Z'1.3 TeV	Impact on sensitivity [%]
Luminosity	2.5	3.7	0.4
PDF uncertainty	3.1	1.0	0.2
$t\bar{t}$ normalization	4.9	$--$	0.7
$t\bar{t}$ ISR, FSR	6.3	$--$	0.7
$t\bar{t}$ fragmentation & parton shower	3.4	$--$	0.9
$t\bar{t}$ generator dependence	2.8	$--$	2.2
$W+$ jets normalization	4.3	$--$	1.4
$W+$ jets shape	$--$	$--$	0.1
Multijets normalization	2.1	$--$	0.2
Multijets shape	$--$	$--$	1.1
$Z+$ jets normalization	2.0	$--$	0.5
Jet energy and mass scale	6.7	2.0	5.2
Jet energy and mass resolution	4.7	4.0	1.2
Electron ID and reconstruction	1.1	1.3	1.0
Muon ID and reconstruction	2.2	2.1	4.8

In the first two columns the relative impact (in percent) is shown on the total expected background yield (nominally 1835 events) and on the number of selected signal events (a Z' with a mass 1.3 TeV is chosen as the benchmark). The final column lists the relative variation of the expected limit on cross section times branching fraction of this benchmark if the corresponding systematic effect is ignored.