

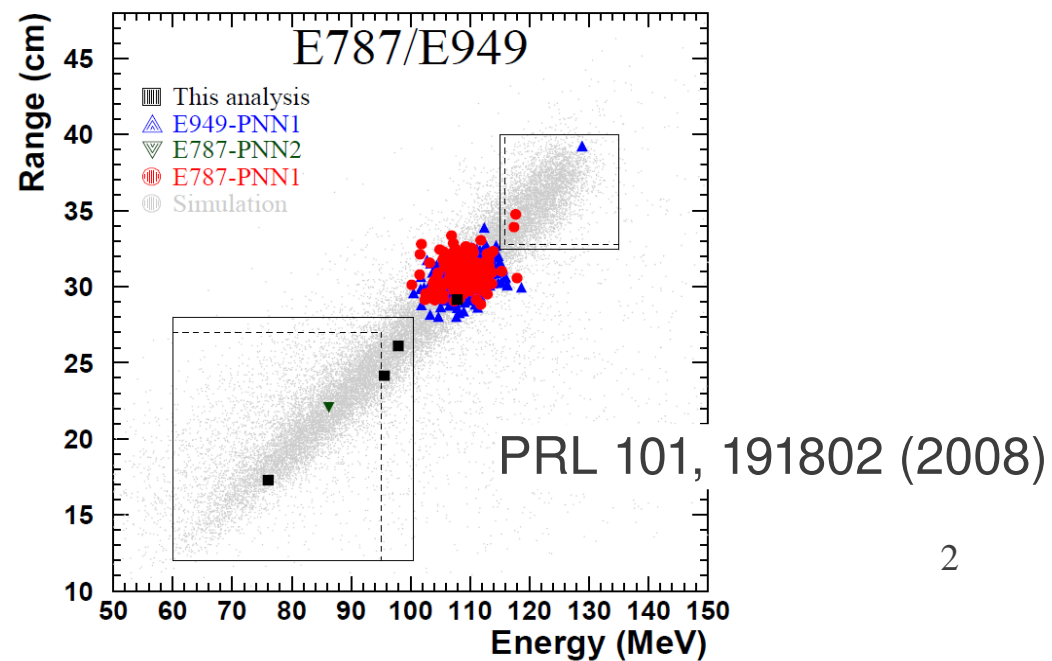
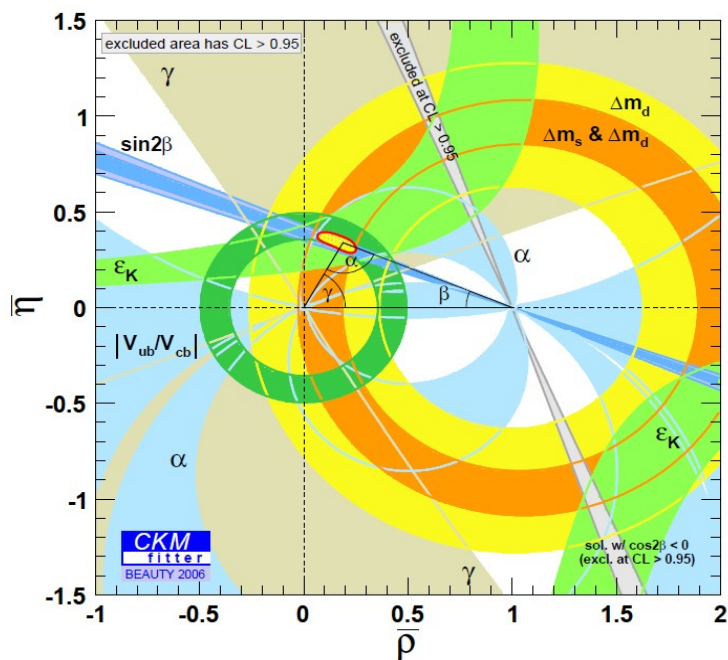


Towards a boosted $VH \rightarrow b\bar{b}$ measurement with ATLAS

LPNHE-Paris seminar
19 January 2012

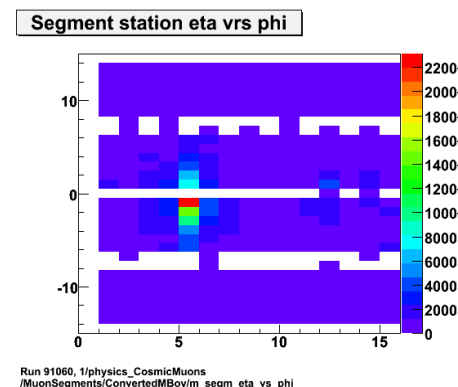
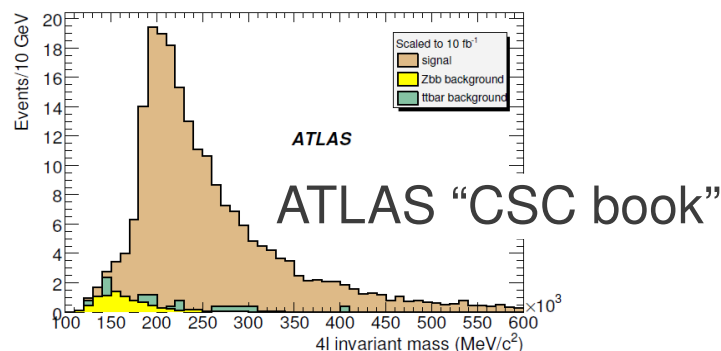
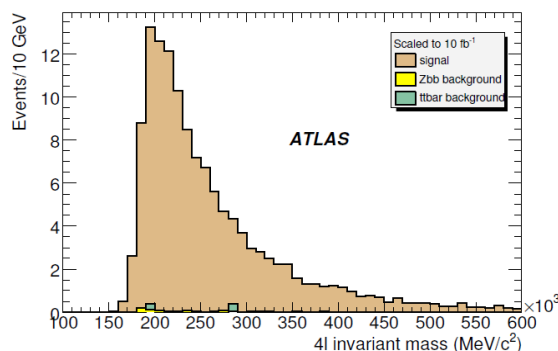
Ilektra Christidi
(CNRS CR1/CR2 candidate)

- 1996-2001: undergraduate, Aristotle University of Thessaloniki, Greece. Graduating grade: 7.9/10, thesis: “Test setup for MDT chambers for ATLAS in the X5/GIF area at CERN”
- 2001-2006: graduate studies, SUNY Stony Brook, NY, USA.
 - MA in 2003, grade 3.82/4, with KOPIO experiment, BNL.
 - PhD in 2006, thesis: “Search for the rare decay $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ with $p_{\pi^+} < 199 \text{ MeV}/c$ ”, with E949 experiment, BNL.



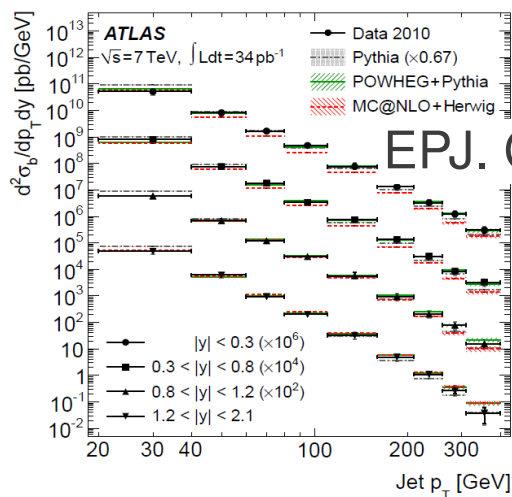
2007-2009: Marie Curie Experienced Researcher, Aristotle University of Thessaloniki, Greece, with ATLAS.

- $ZZ^{(*)} \rightarrow 4l$ sensitivity studies in MC, Muon Spectrometer commissioning and DQM
- Single lepton efficiencies from data, experimental signatures to differentiate between BSM scenarios, data-driven method to estimate Z+jets background to ZZ using same-sign leptons.

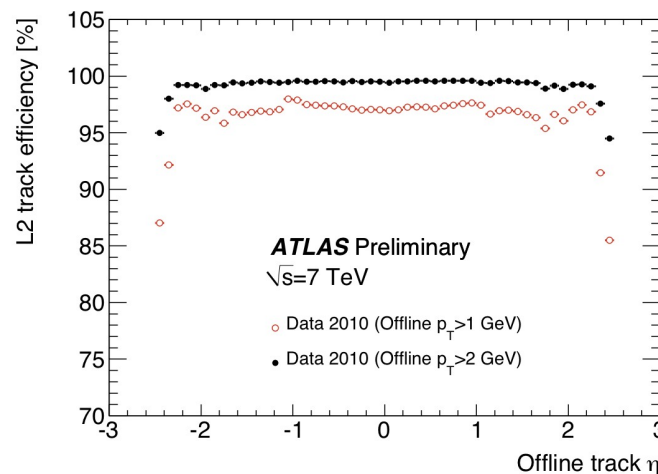


2009-2012: postdoctoral researcher partly funded by Marie Curie, UCL, UK, with ATLAS.

- B-jet and bb-dijet cross-section measurement, online tracking for HLT
- Currently: b-tagging in boosted jets, in the boosted $Z \rightarrow bb$ and $VH \rightarrow bb$ framework

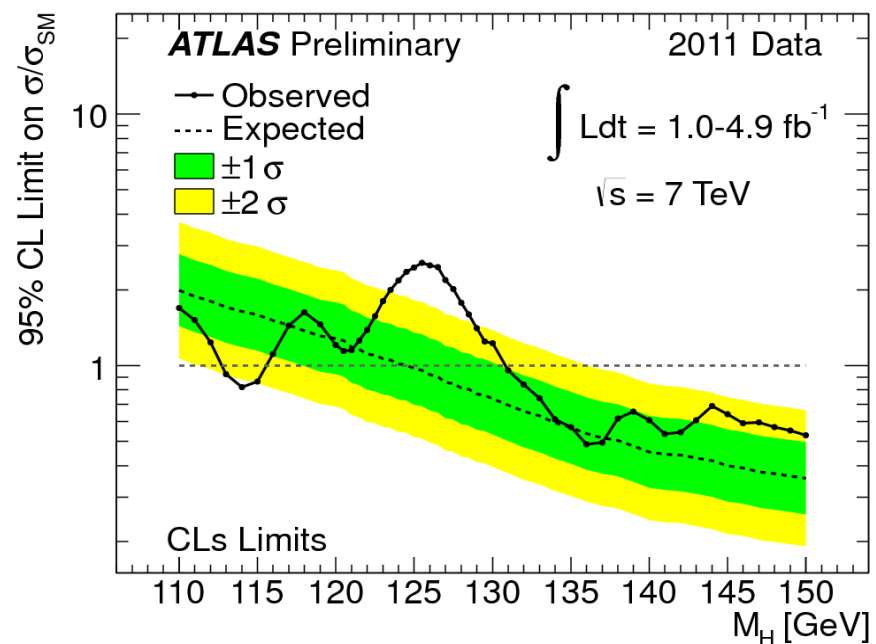
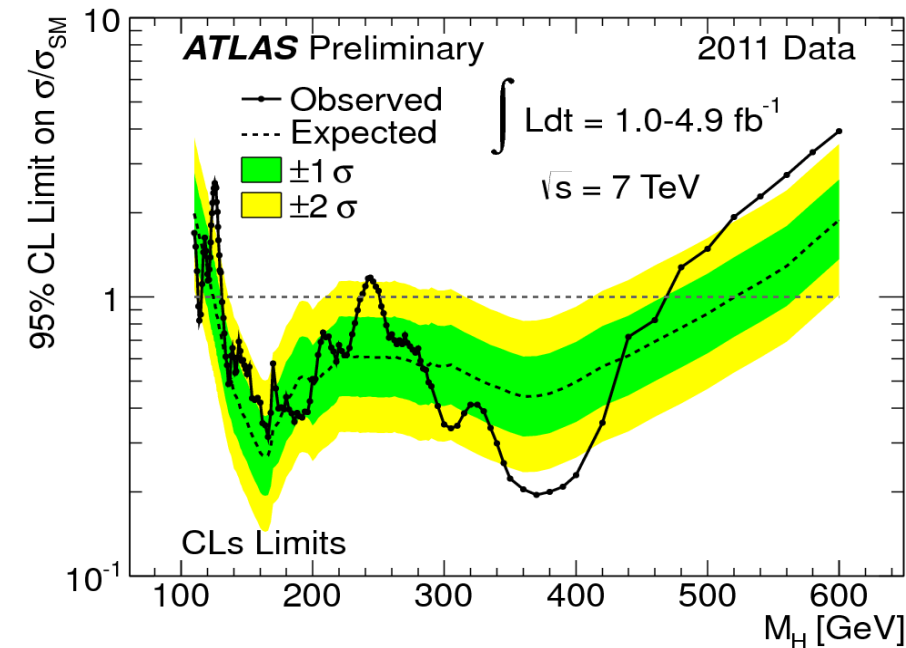


EPJ. C 71, 1846 (2011)



- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - proof of principle of a basic b-jet measurement (EPJ C 71, 1846 (2011))
 - Jet substructure measurements
 - need to understand substructure in order to use it (ATLAS-CONF-2011-073)
 - Un-boosted (inclusive) $VH \rightarrow bb$
 - many common elements (ATLAS-CONF-2011-103)
- Conclusions

- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions



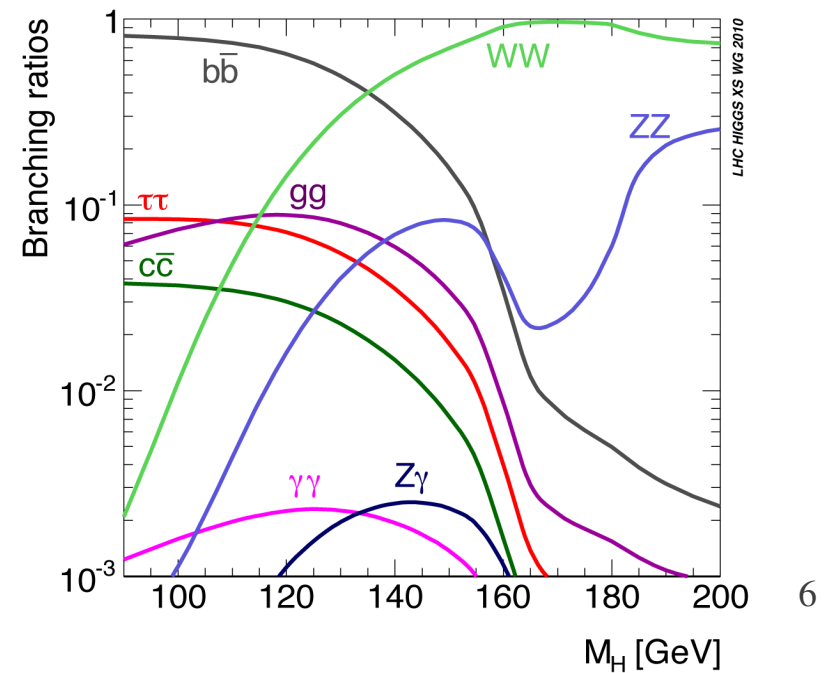
Current allowed regions (ATLAS only):

$$115.5 \text{ GeV} < m_H < 131 \text{ GeV}$$

$$237 \text{ GeV} < m_H < 251 \text{ GeV}$$

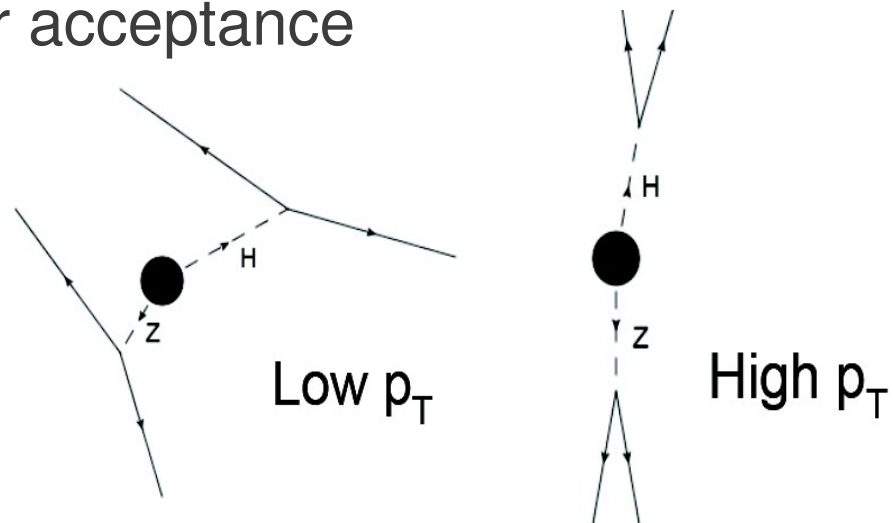
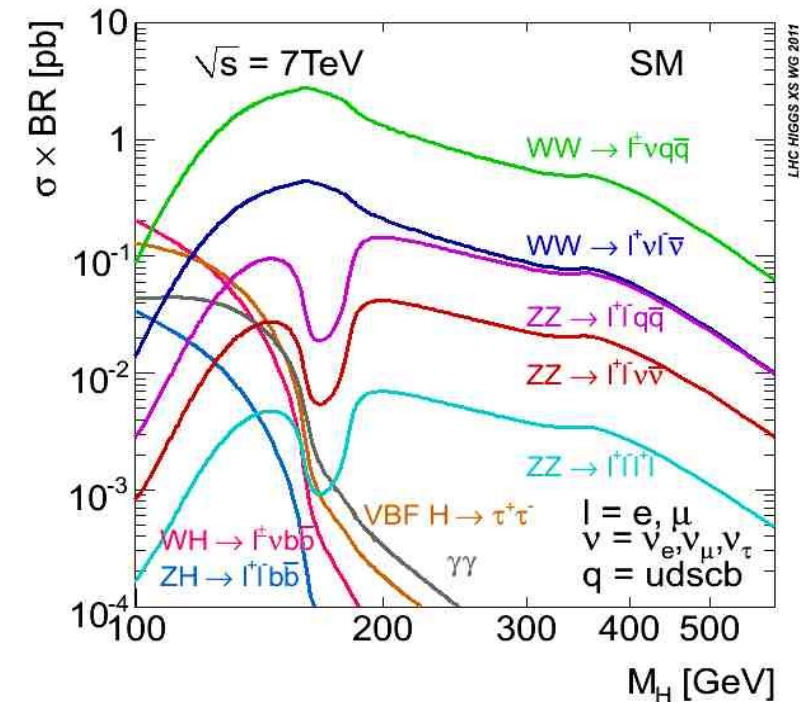
$$m_H > 468 \text{ GeV}$$

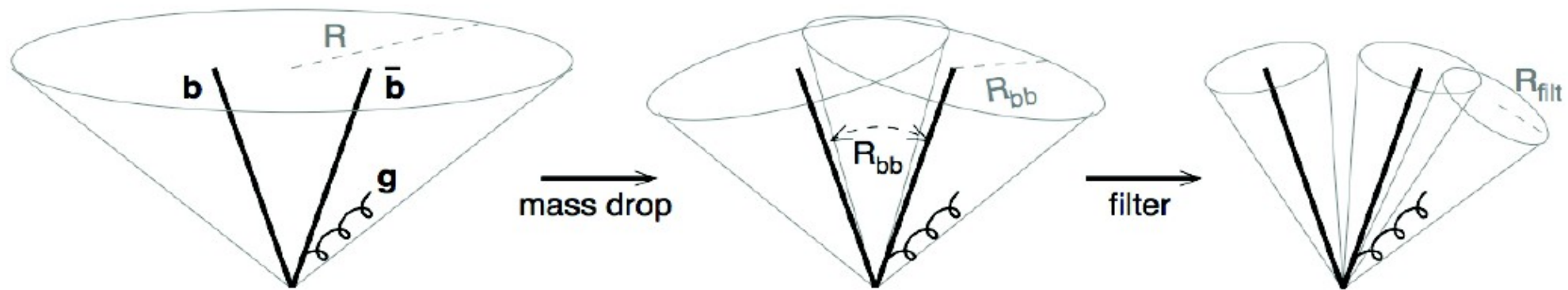
In the low-mass region, $H \rightarrow b\bar{b}$ is the dominant decay channel



Why boosted $VH \rightarrow bb$?

- $gg \rightarrow H \rightarrow bb$ overwhelmed by backgrounds, therefore VH ($V=W$ or Z) associated production more sensitive
- Still, $VH \rightarrow bb$ quite background-prone.
- New idea: use only highly boosted H/V , which costs $\sim 95\%$ signal acceptance, but reduces backgrounds drastically (PRL 100, 242001 (2008) J. Butterworth et al.)
- Higgs decay products in one, fat b-jet. All final particles within detector acceptance



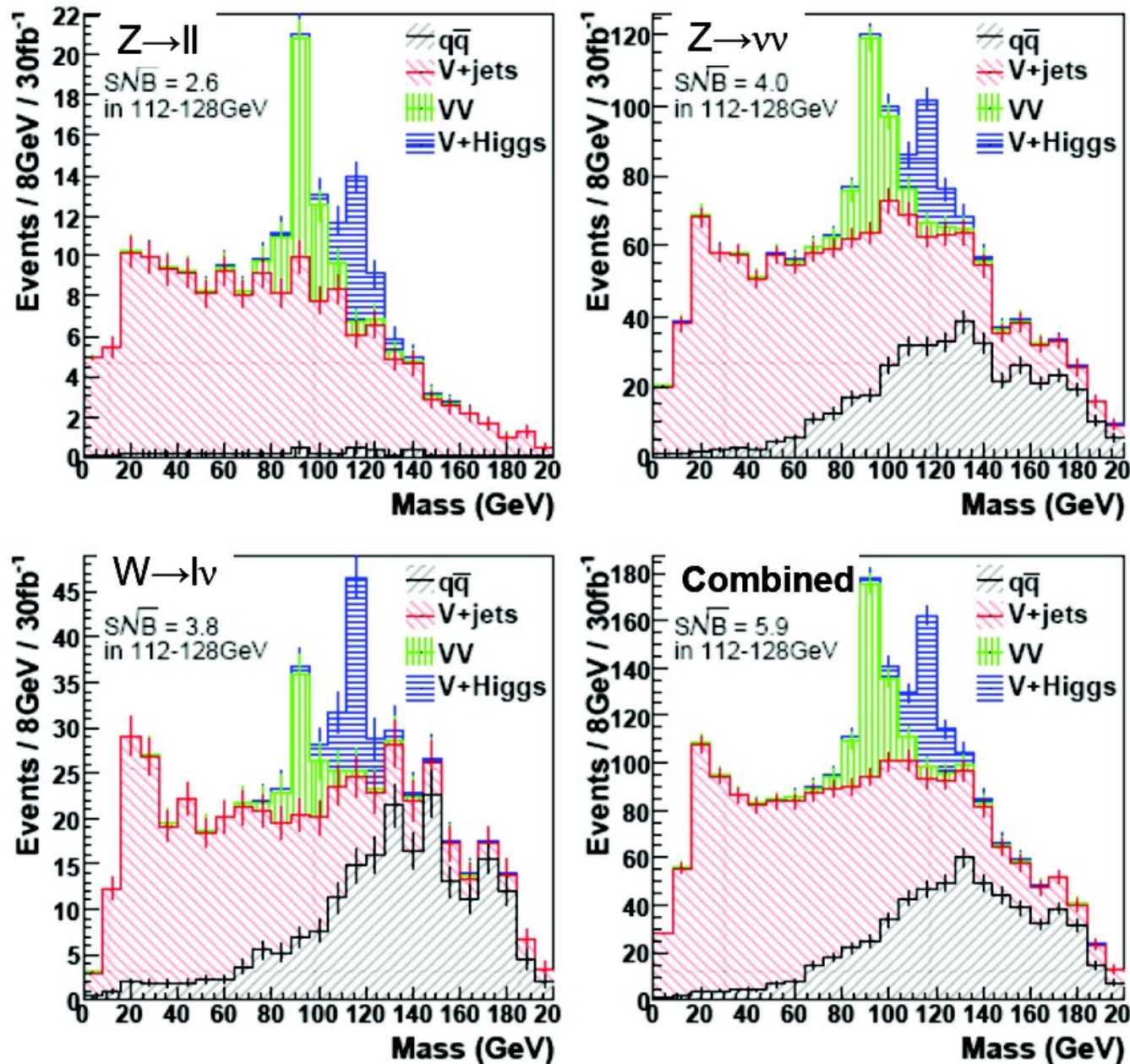


Use the angle-ordered Cambridge-Aachen (C-A) jet algorithm, which keeps track of clustering steps.

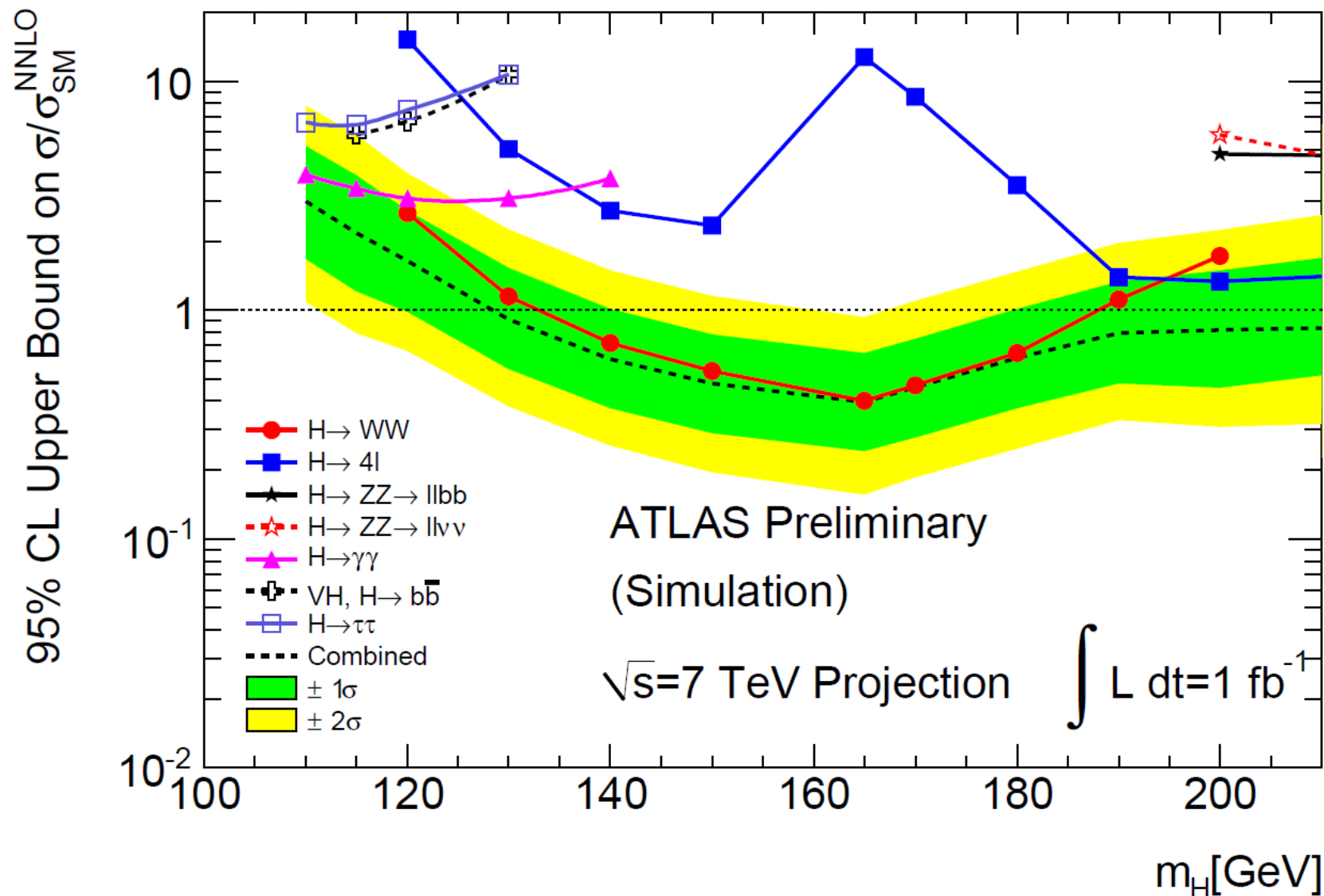
- Clustering can be undone one step at a time with an angular distance $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$ of R_{bb} , until a large drop in mass is observed
- Check this splitting is not too asymmetric
- Recluster remaining constituents with smaller $R_{filt} = \min(0.3, R_{bb}/2)$ and keep up to 3 constituents to reconstruct the Higgs mass

In the following, consider C-A jets with $R=1.2$

From PRL 100, 242001 (2008) (particle level simulation), 30 fb^{-1}

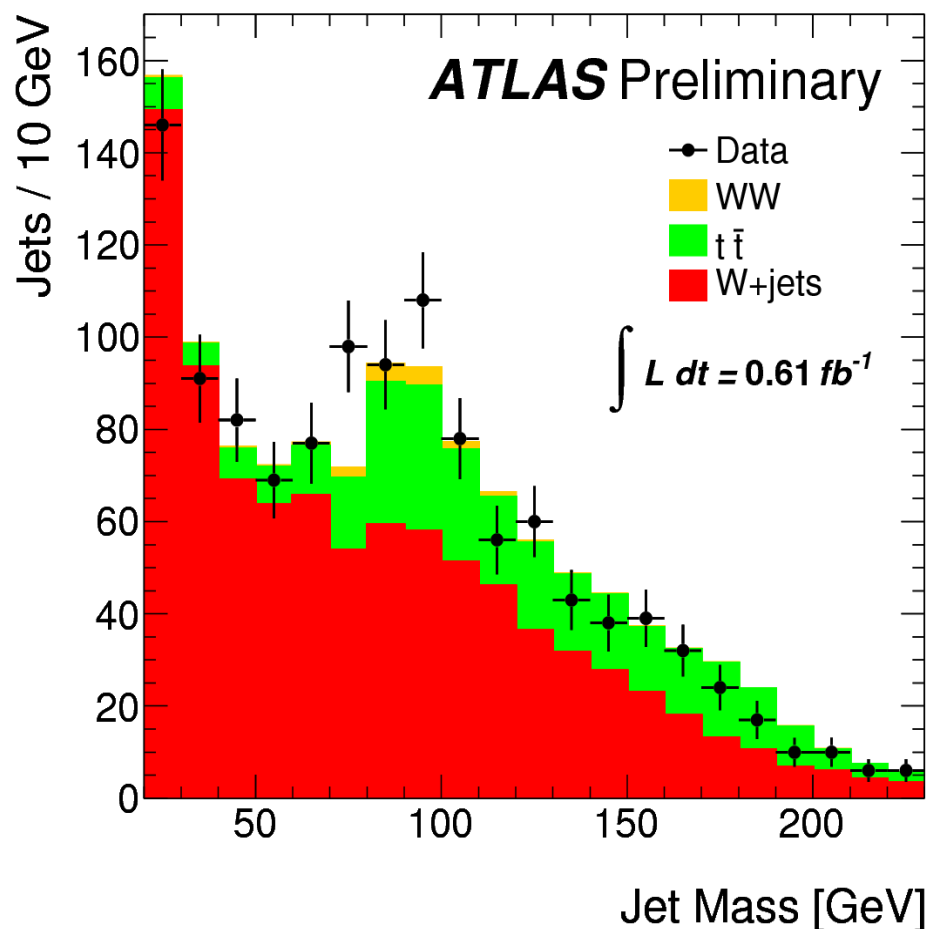


From ATLAS-PUB-2010-015 (full ATLAS detector simulation), 1 fb^{-1}

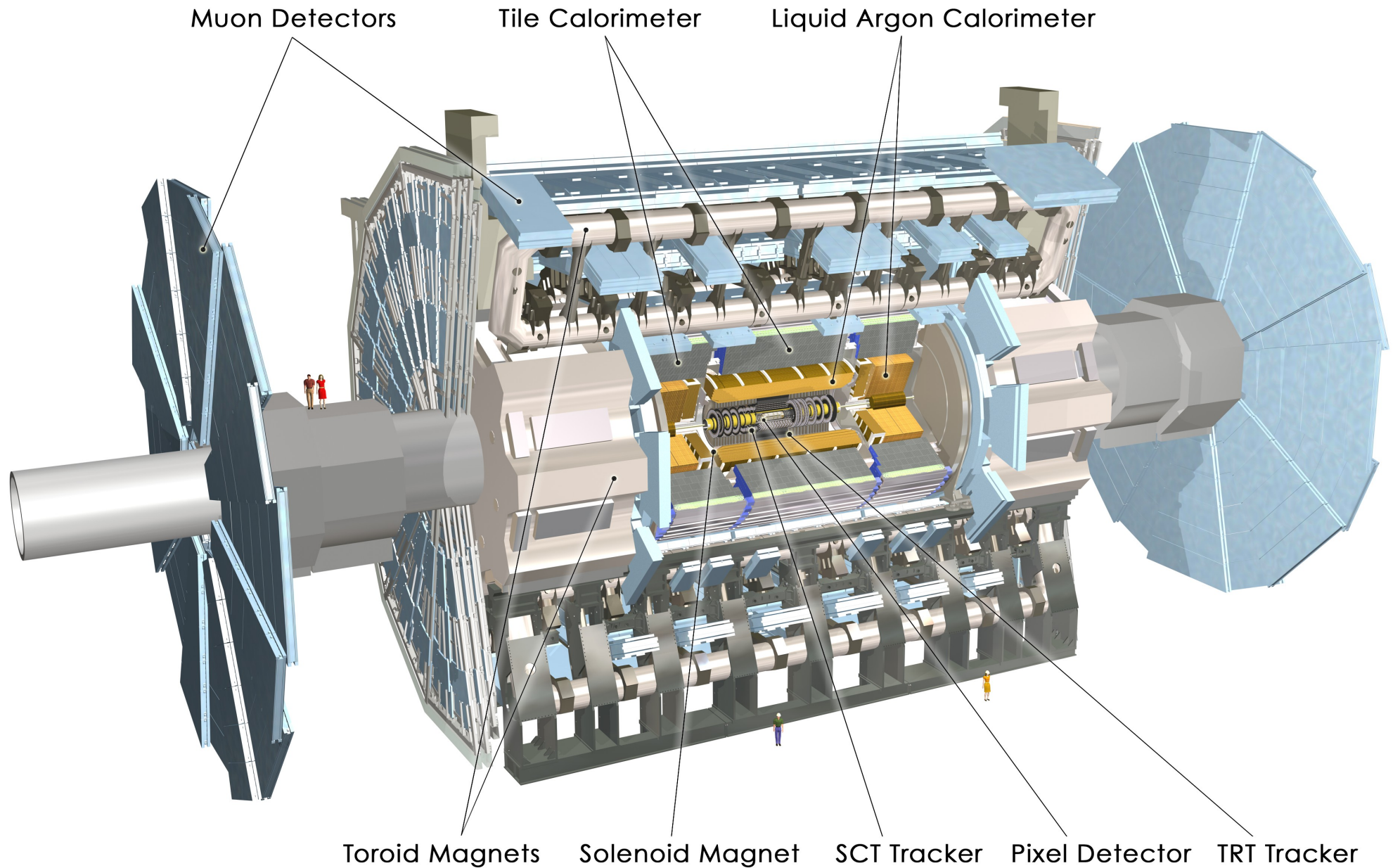


Very encouraging first result from summer 2011:

Split and filtered C-A 1.2 jet mass, in association with a $W \rightarrow lv$ with $p_T > 200 \text{ GeV}$. No b-tagging applied.



- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions

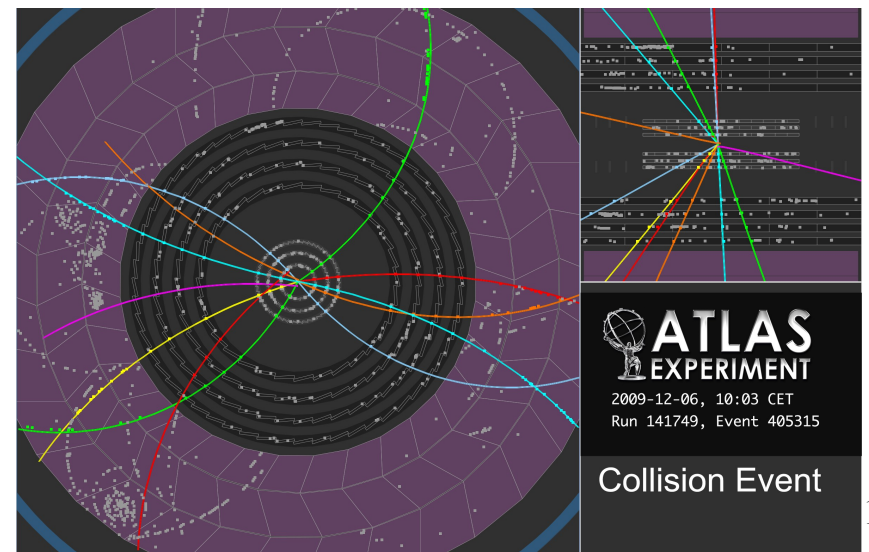
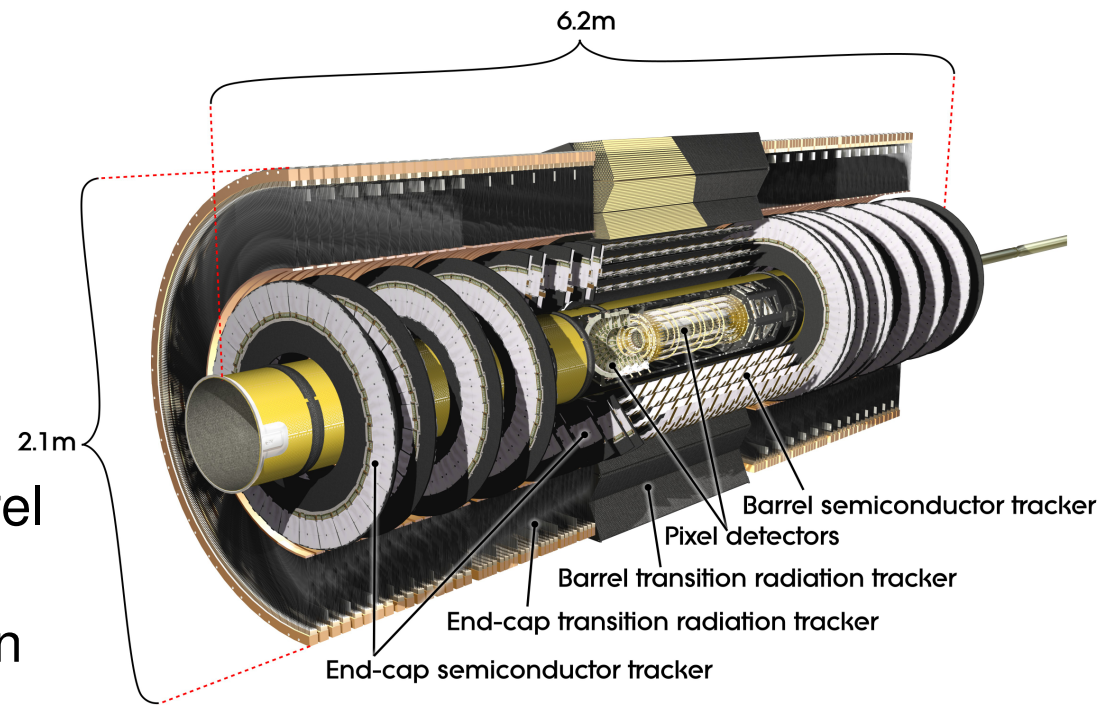


Detectors of 3 different technologies for optimal position detection and track reconstruction:

- silicon **pixels** (3 layers)
- silicon strips – **SCTs** (4 layers in Barrel and 9 in EC). Each layer contains 2 sub-layers in an angle, for 3D position information
- Transition Radiation Tracker - **TRT**

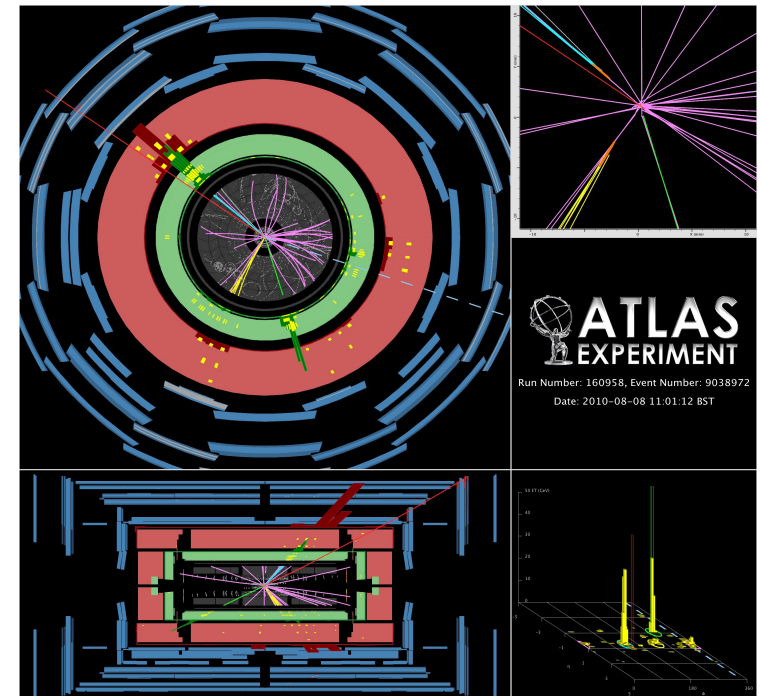
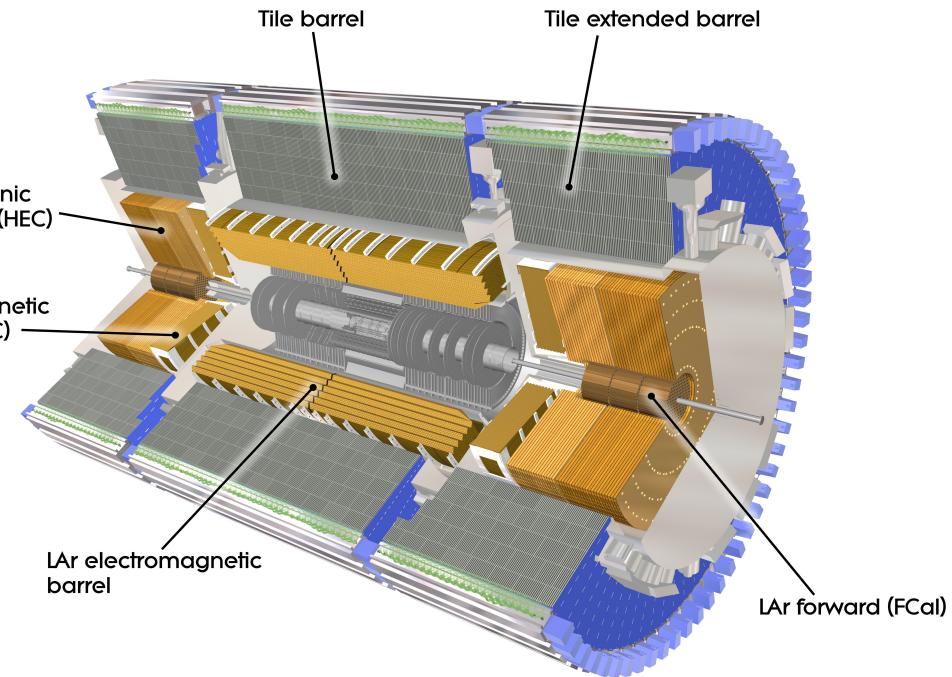
The whole structure is immersed in a 2T solenoid magnetic field.

→ **Track-based *b*-tagging happens here**



Detectors of 2 different technologies for electromagnetic and hadronic calorimetry:

- Pb (barrel) or copper (EC) absorber-Liquid Argon (**LAr**) for EM in barrel and both EM and HAD in EC and Forward
- Iron absorber-scintillator **Tile** for HAD in barrel

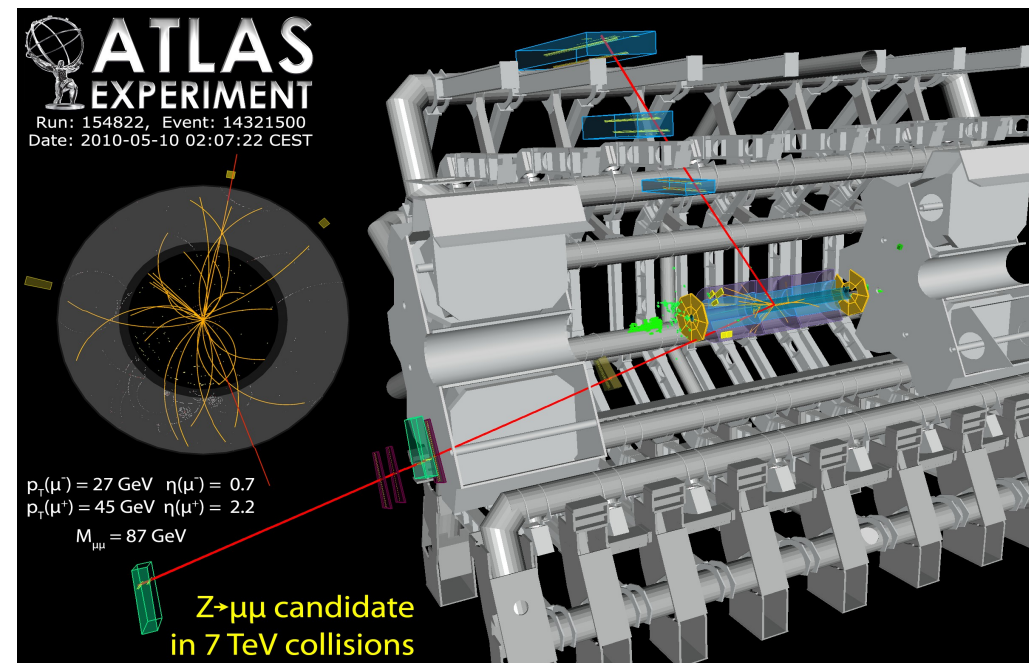
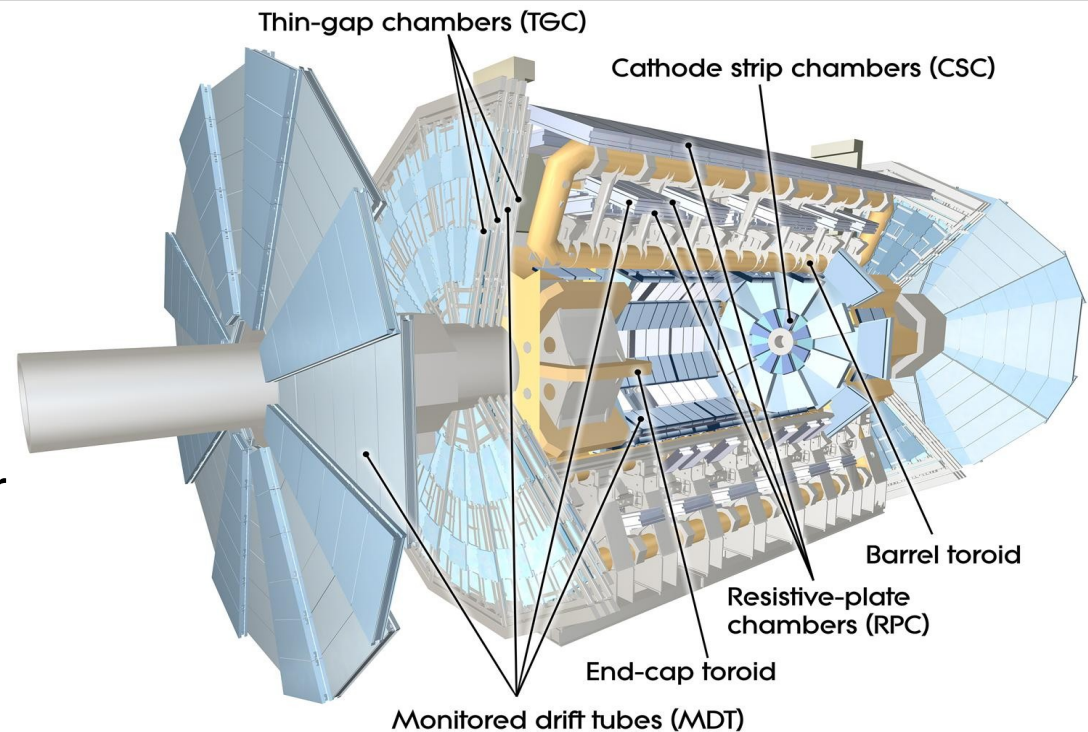


→ *Electrons, photons and jets are reconstructed here*

Chambers of 4 different technologies for optimal muon reconstruction and trigger:

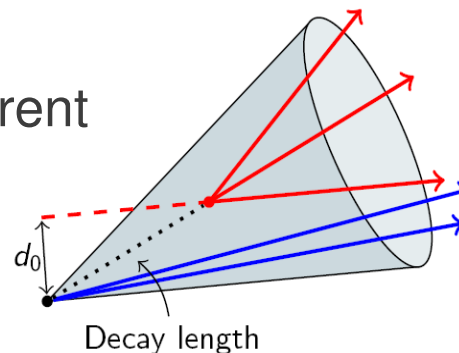
- Monitored Drift Tubes (**MDTs**) for precision position measurement
- Resistive Plate Chambers (**RPCs**) for triggering in the Barrel
- Thin Gap Chambers (**TGCs**) for triggering in the ECs
- Cathode Strip Chambers (**CSCs**) for tracking in the busier, high- η region

The whole structure is immersed in a 0.5T toroid magnetic field.



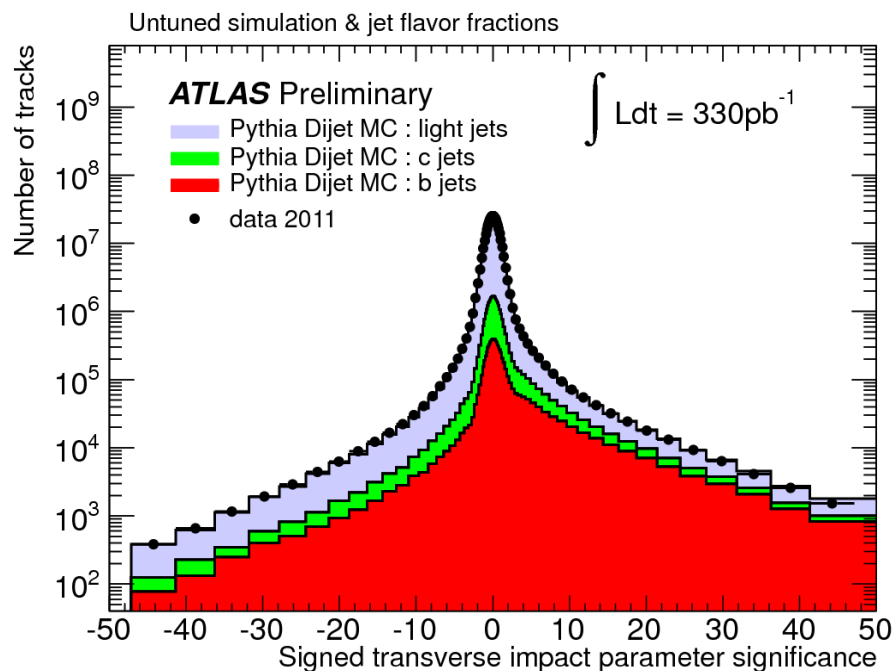
→ **Muons are reconstructed in the MS and ID**

Tracks from the **b-decay** have different topology than **prompt** ones



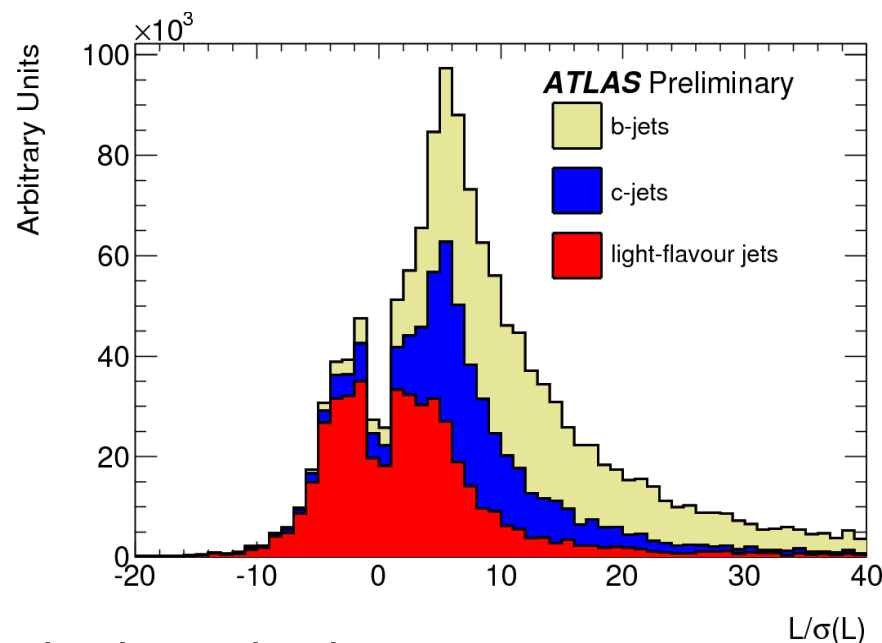
Impact Parameter-based algorithms:

exploit the fact that b-decay tracks are not compatible with Primary Vertex $\rightarrow$ higher impact parameter



Secondary Vertex-based algorithms:

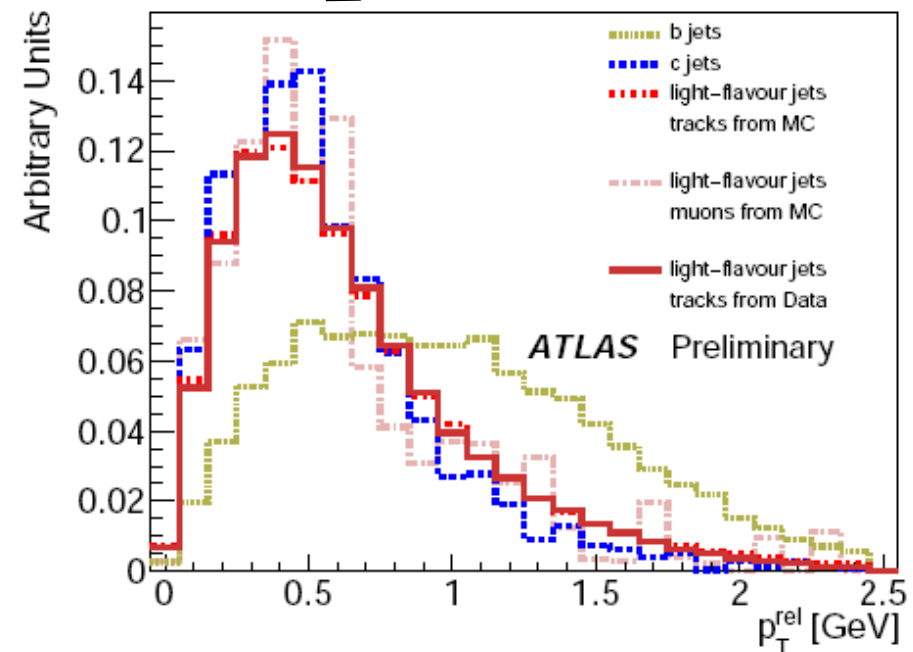
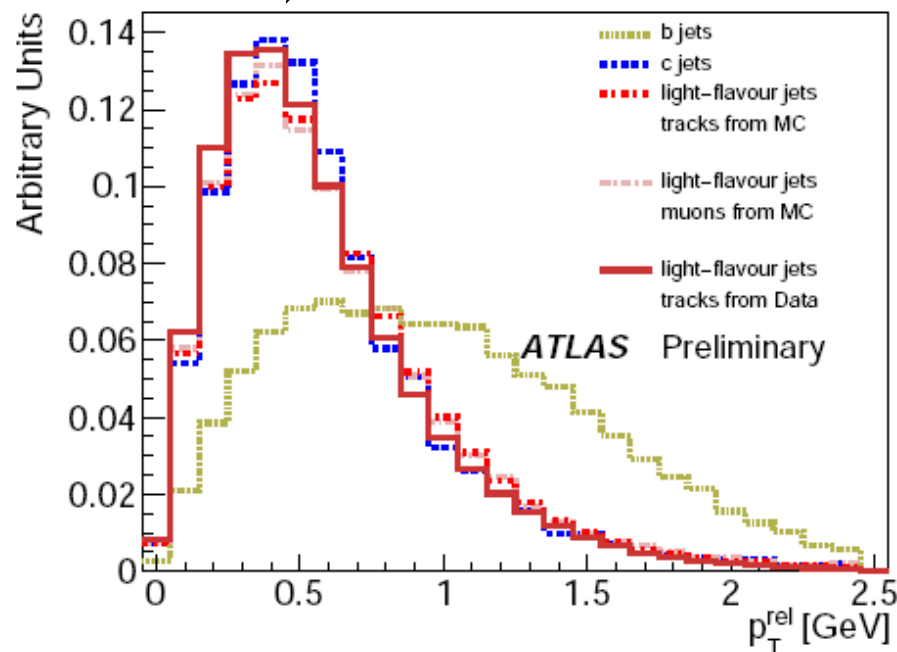
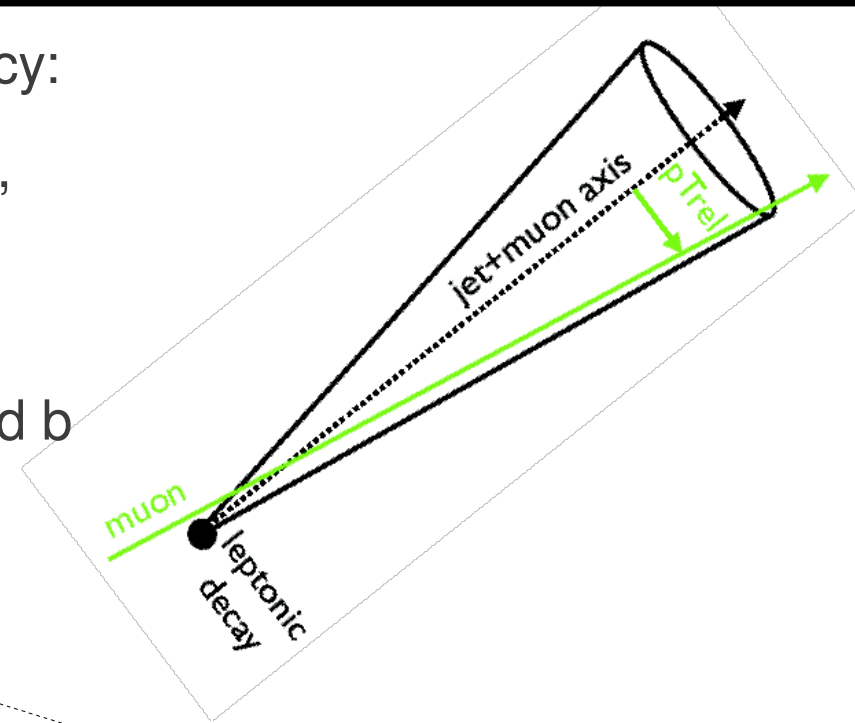
try to reconstruct b-decay vertex. SV0 starts from two-track vertices and the signed decay length significance is its b-tag weight of the jet



Combined algorithms: use information from both methods

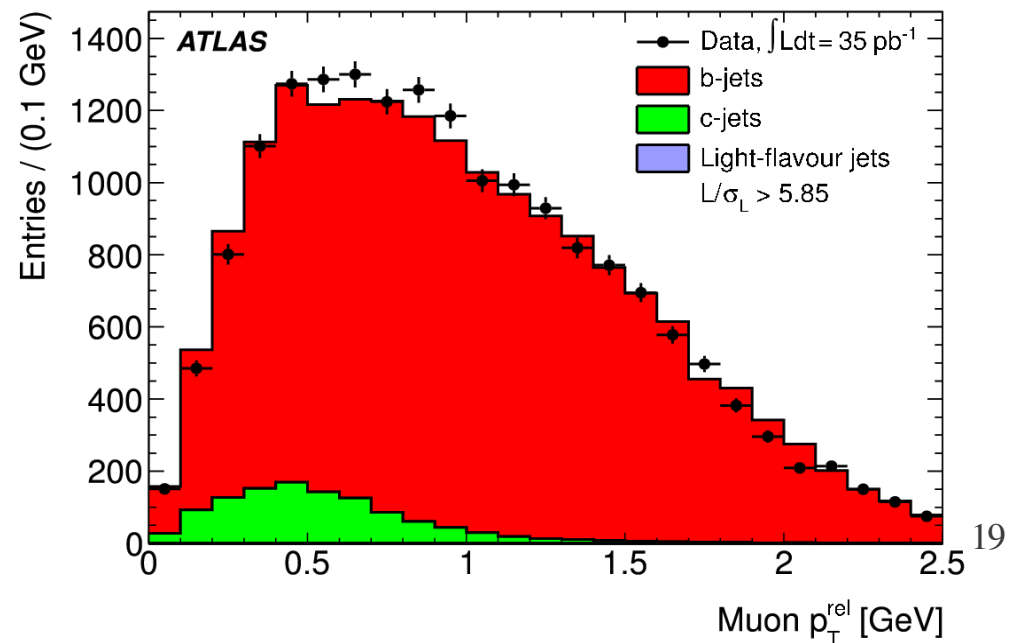
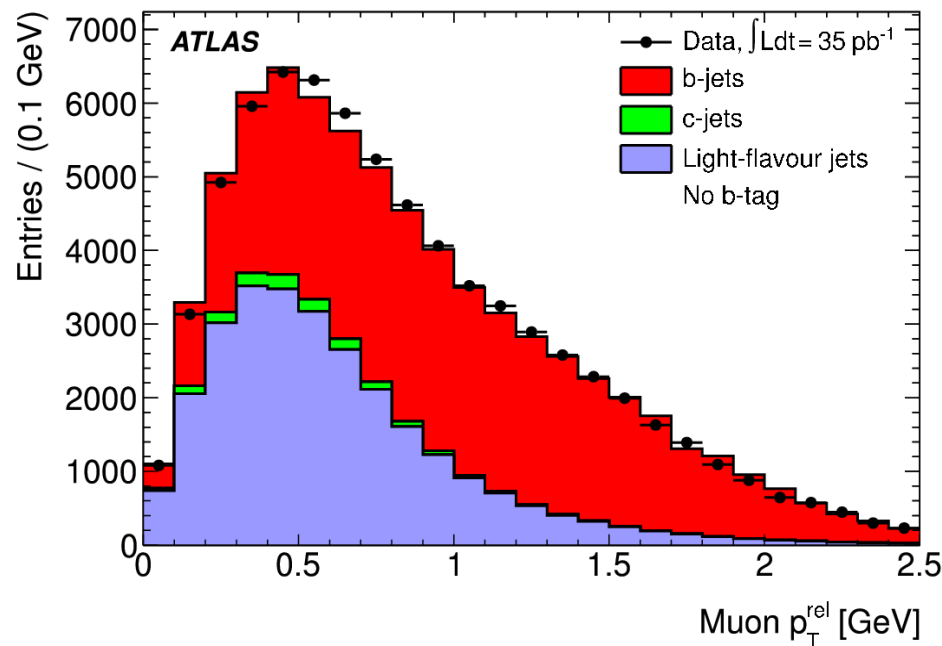
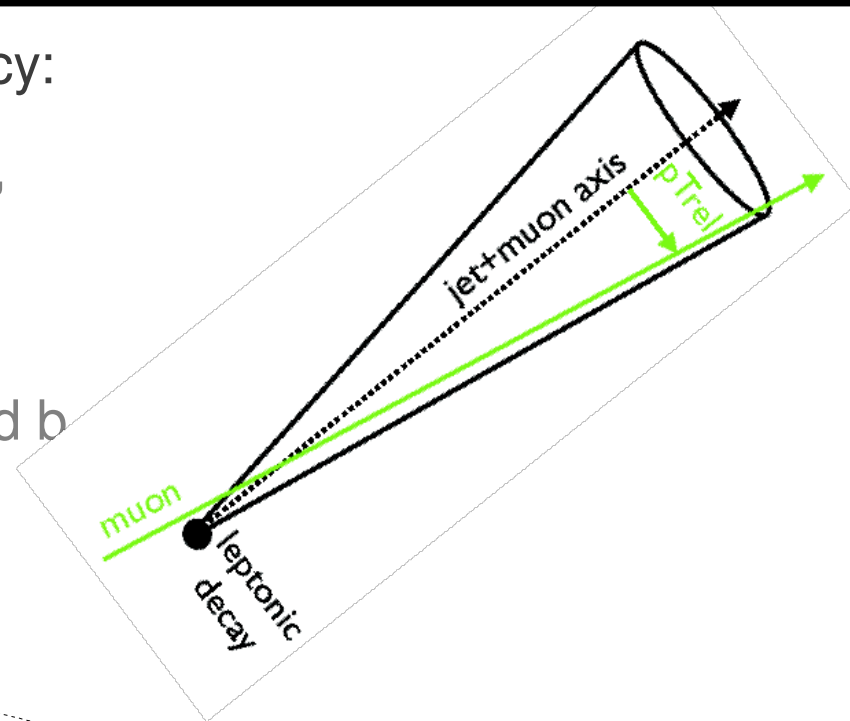
Data- driven method to calculate b-tagging efficiency:

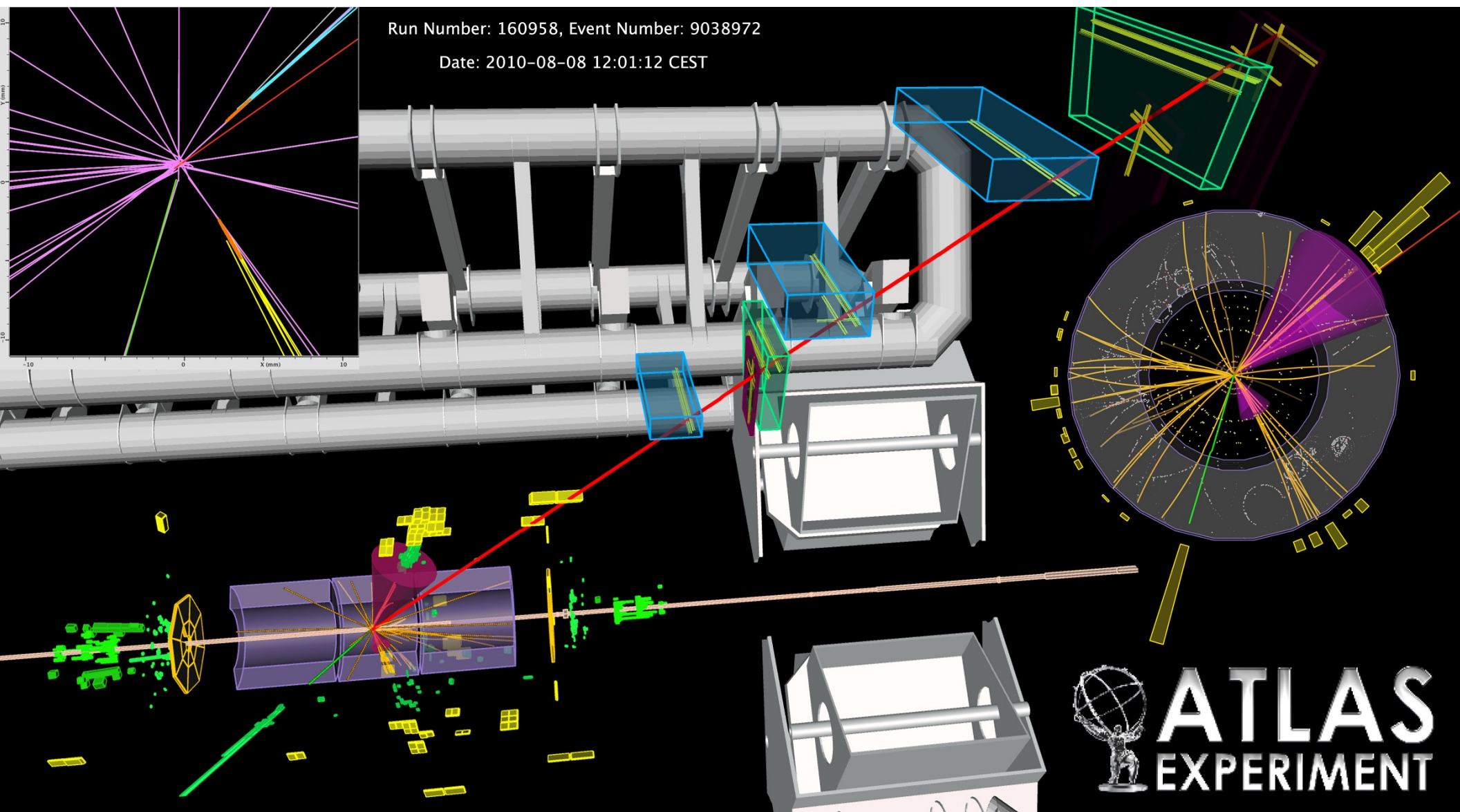
- Select jets with muon in them $\rightarrow$ enhanced in b, independent of track-based b-tagging
- Get template of relative p_T of the muon with respect to the jet axis (p_T^{rel}) for light (data), c and b (MC) jets, before and after SV0 cut
- Fit the data to them before and after SV0 cut $\rightarrow$ ratio of number of b-jets is the efficiency

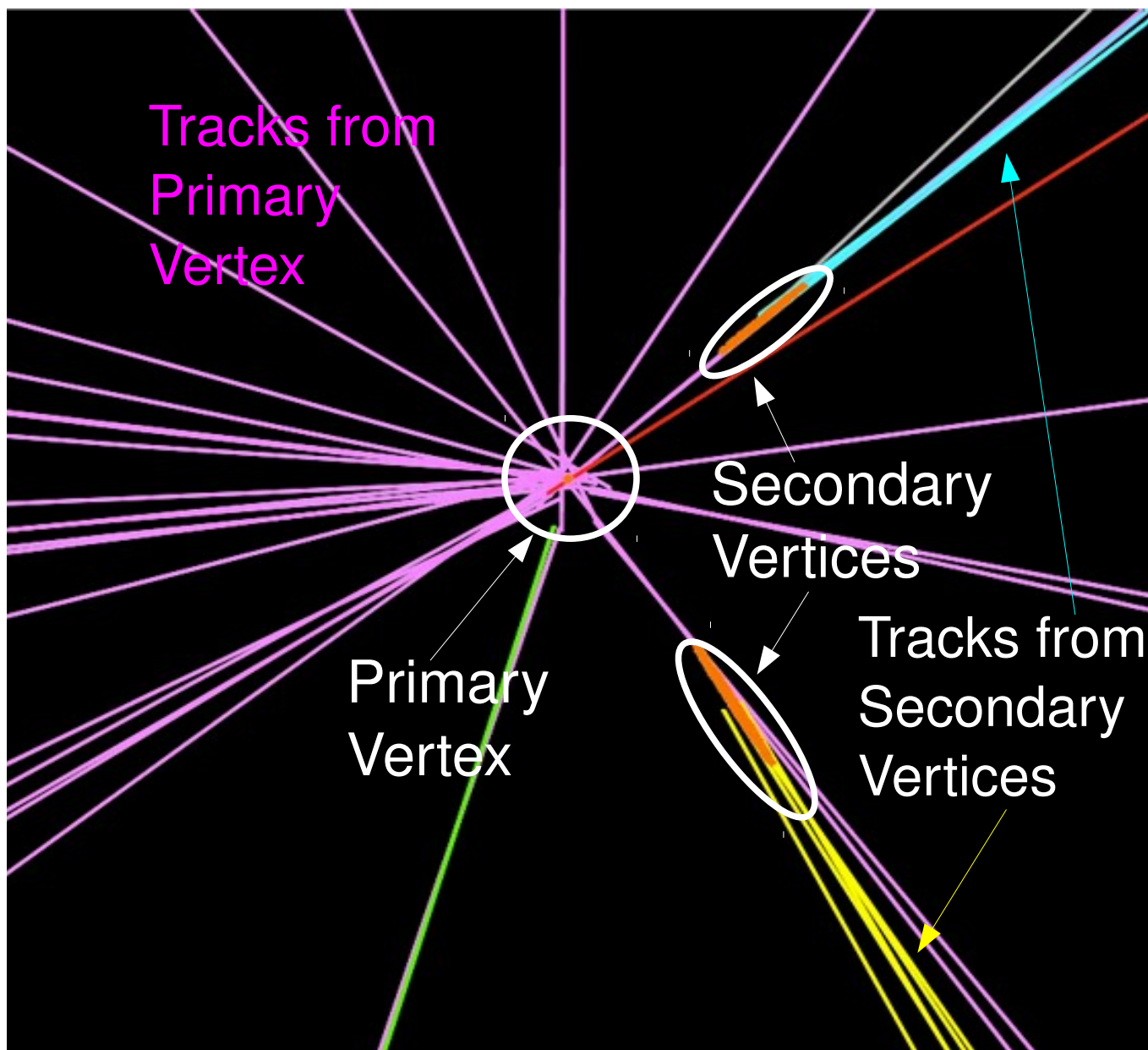


Data- driven method to calculate b-tagging efficiency:

- Select jets with muon in them $\rightarrow$ enhanced in b, independent of track-based b-tagging
- Get template of relative p_T of the muon with respect to the jet axis (p_T^{rel}) for light (data), c and b (MC) jets, before and after SV0 cut
- Fit the data to them before and after SV0 cut $\rightarrow$ ratio of number of b-jets is the efficiency







- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions

- Motivation: Constrain b-component of proton PDF, compare with QCD MC models
- Use SV b-tagging to identify b-jets, then apply efficiency correction
- Get b-fraction in final sample from binned likelihood fit of SV mass with MC templates
- Results presented with $\sim 34 \text{ pb}^{-1}$ of data, using single-jet and minimum bias triggers

Number of SV0-tagged jets

Fraction of them that are b-jets, found from template fit

$$\frac{d^2 \sigma_b}{dp_T dy} = \frac{1}{\Delta p_T \Delta y} \frac{N_b \cdot \text{frac}_b}{\underbrace{\epsilon_{trig} \cdot \epsilon_{sel} \cdot \epsilon_{btag}}_{\text{Efficiencies: trigger, jet reconstruction and selection (>99%), b-tagging.}}} \cdot L \times C$$

Efficiencies: trigger, jet reconstruction and selection (>99%), b-tagging.

Unfolding correction

Luminosity for the triggers used

- Motivation: Constrain b-component of proton PDF, compare with QCD MC models
- Use SV b-tagging to identify b-jets, then apply efficiency correction
- Get b-fraction in final sample from binned likelihood fit of SV mass with MC templates
- Results presented with $\sim 34 \text{ pb}^{-1}$ of data, using single-jet triggers

Number of SV0-tagged di-jets

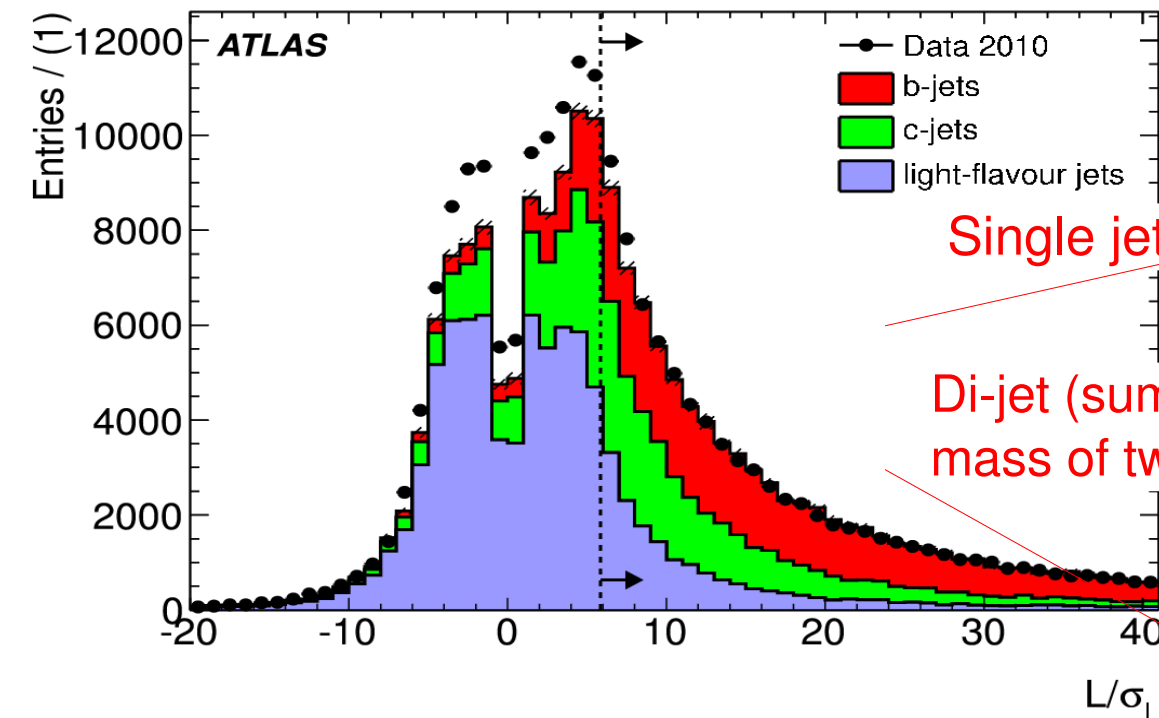
Fraction of them that are b-jets,
found from template fit

$$\frac{d\sigma_b}{dM} = \frac{1}{\Delta M} \frac{N_{b\bar{b}} \cdot \text{frac}_b}{\underbrace{\epsilon_{trig}^{jj} \cdot \epsilon_{sel}^{jj} \cdot \epsilon_{btag}^{jj}} \cdot L} \times C$$

Efficiencies: trigger, jet reconstruction
and selection (>99%), b-tagging.

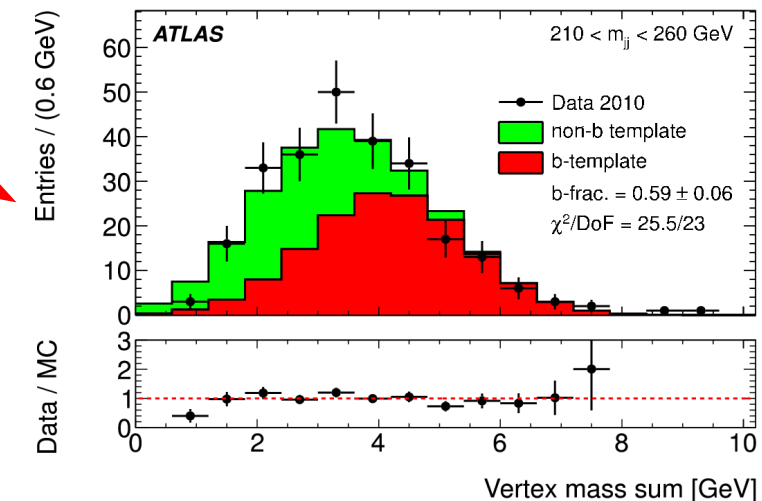
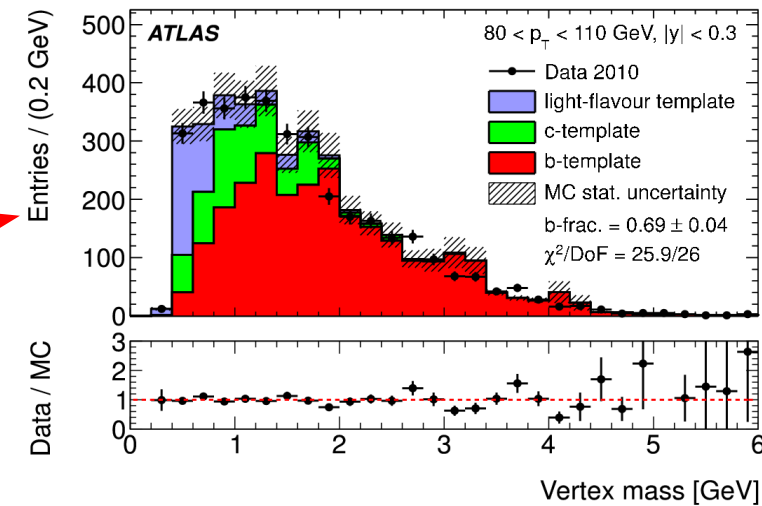
Unfolding correction

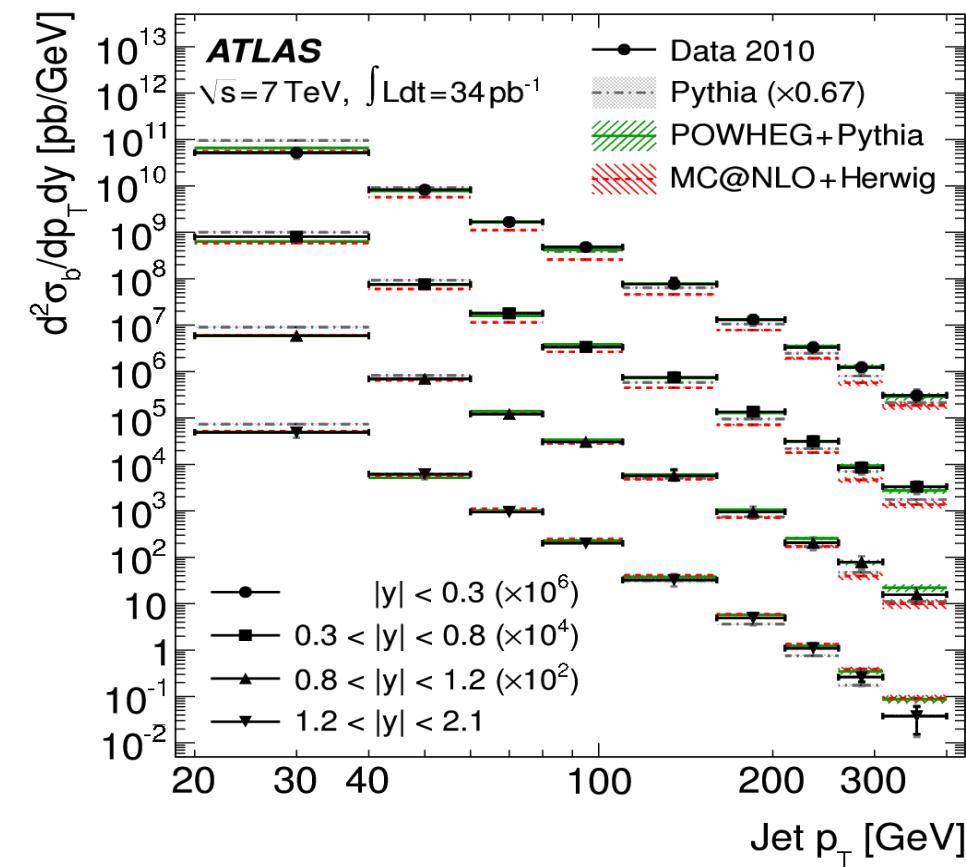
Luminosity for the triggers used



- The SV0 tagger weight is the 3D decay length significance signed w.r.t. the jet axis.
- Cut at 5.85 and use the invariant mass of the tracks in the secondary vertex to separate b- from non-b-jets in a template fit in each p_T - η or di-jet mass bin.
- Take light, charm and b (single-jet) or b and non-b (di-jet) templates from MC
- Log-likelihood fit to data, taking into account the statistical uncertainty on both the data and the MC templates

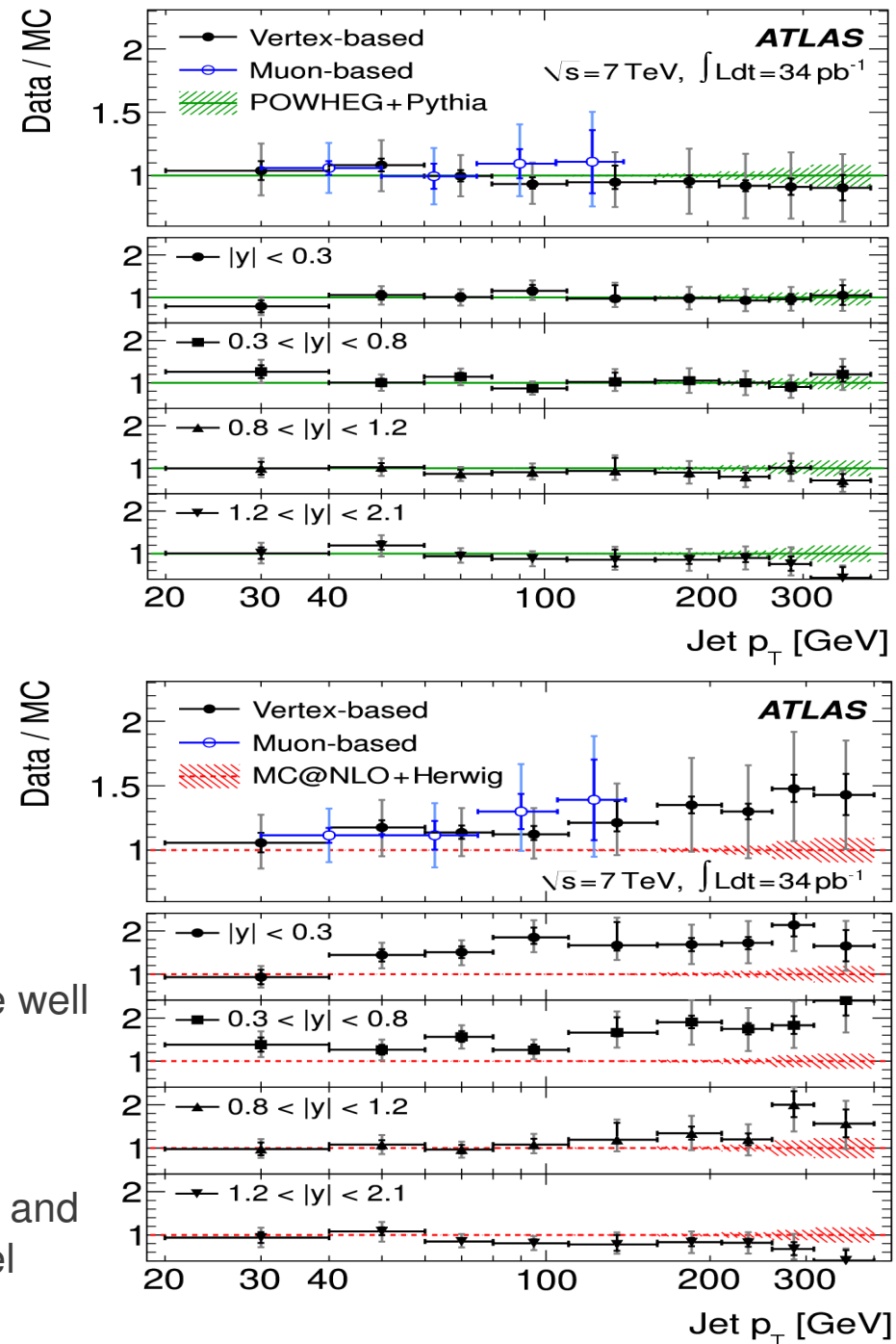
→ *b-jet fraction*

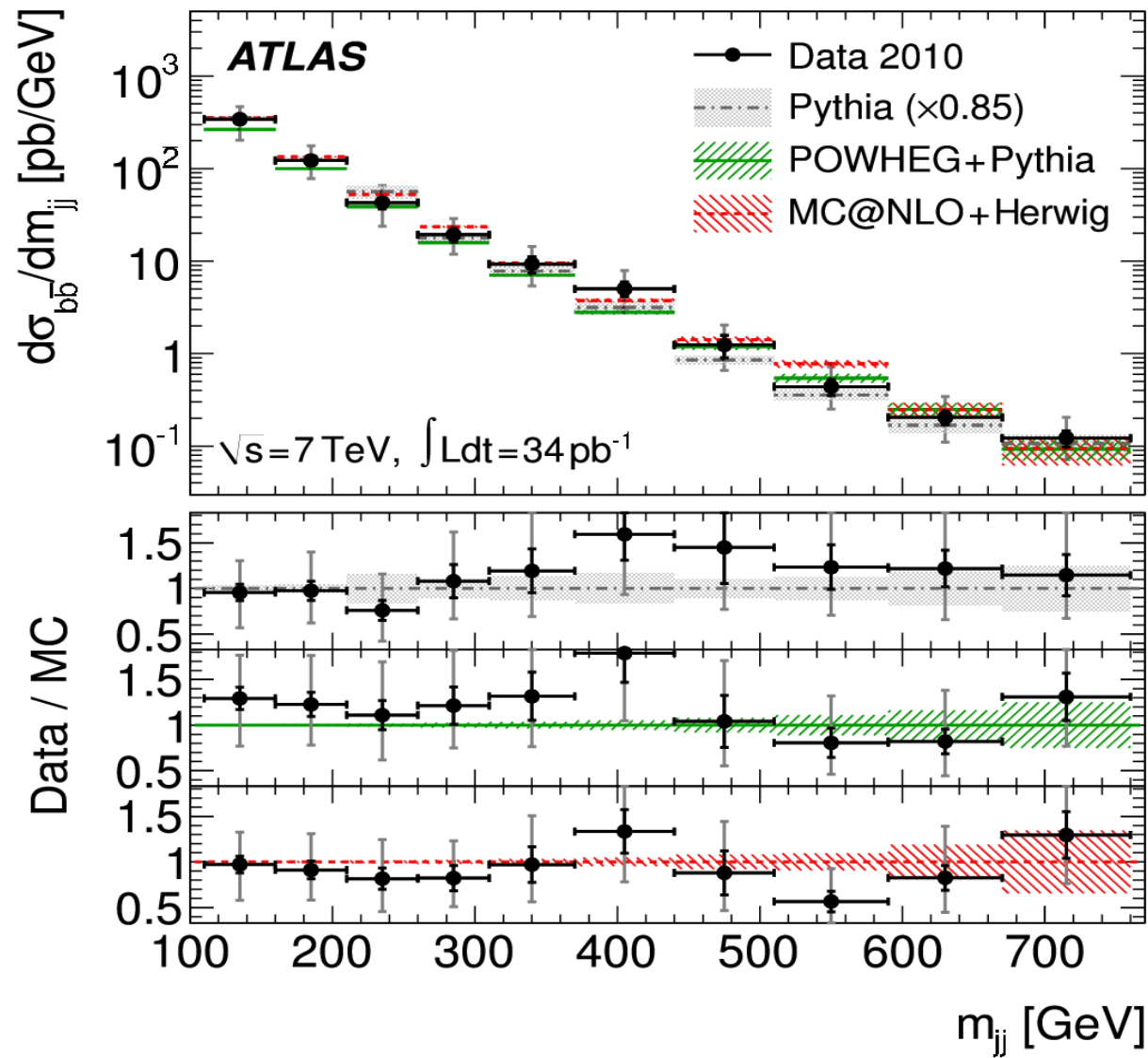




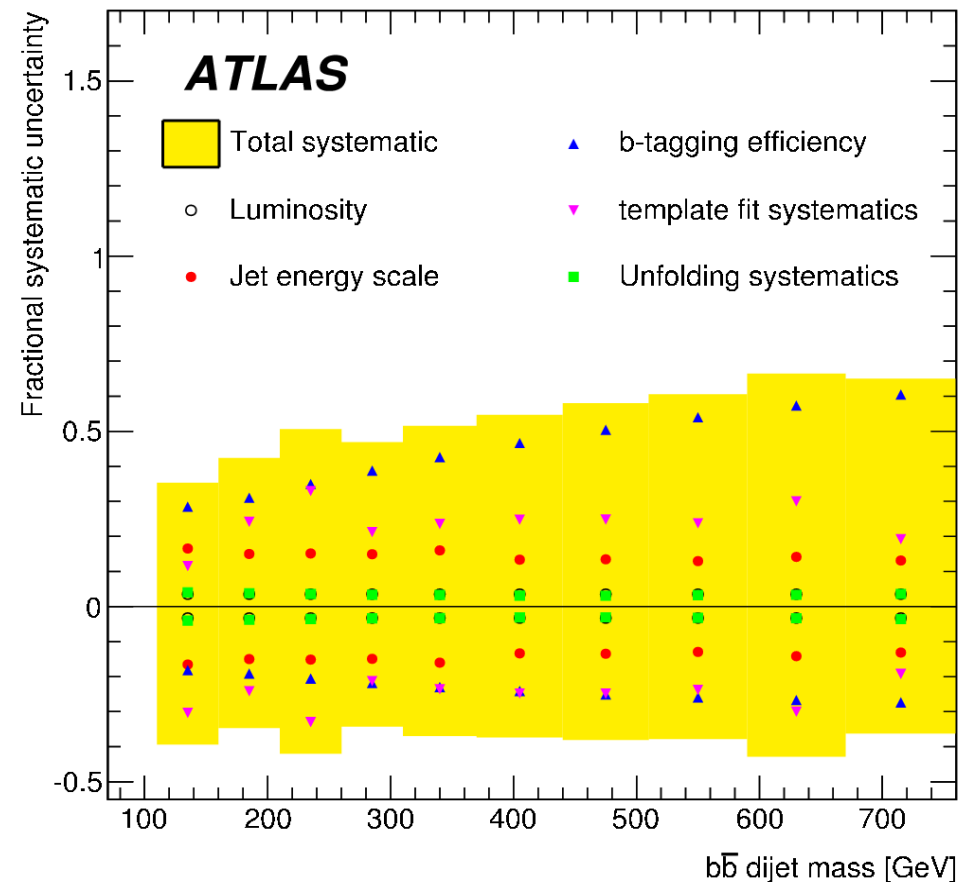
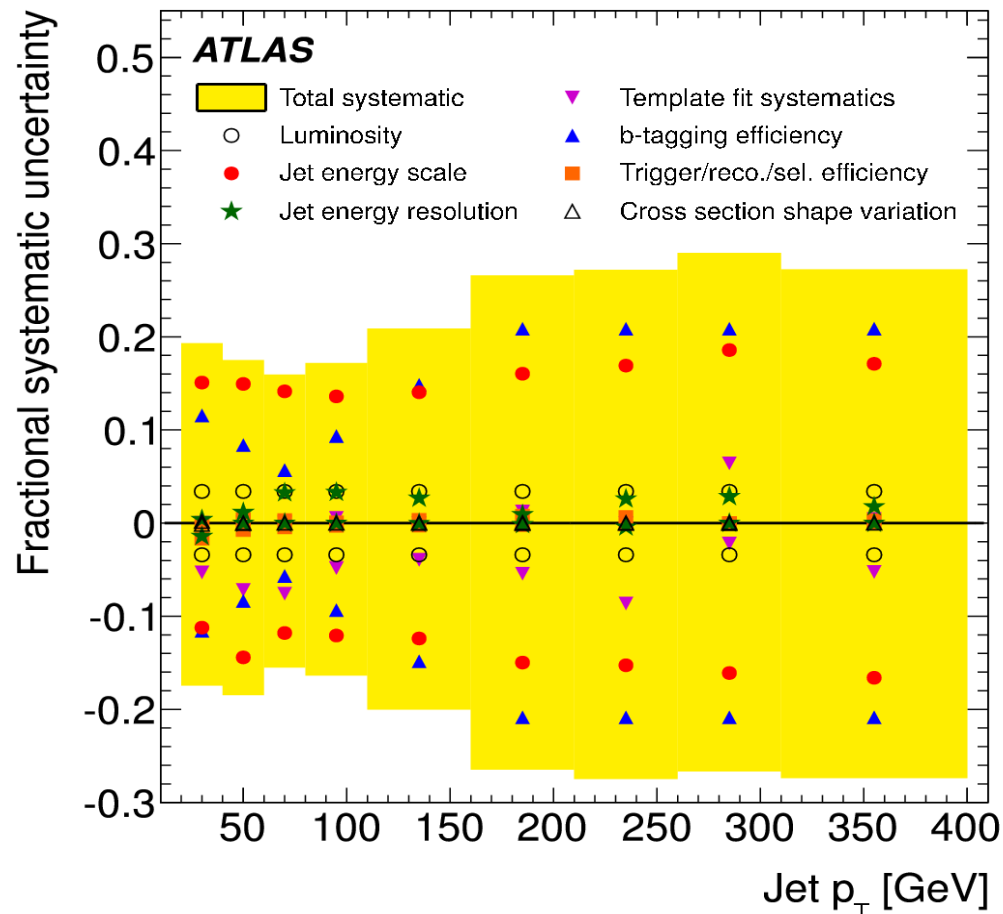
MC comparisons: broad agreement within the systematic uncertainties

- Pythia6 (LL): normalization not expected to agree, shape well described
- POWHEG (NLO) + Pythia: good agreement
- MC@NLO + Herwig: disagreement for central rapidities and higher p_T , probably due to different hadronization model





Good agreement with Pythia, POWHEG and MC@NLO MC



Syst. uncertainty	Inclusive b -jet	$b\bar{b}$ -dijet
Jet energy scale	10–20%	10–20%
b -tagging efficiency	5–20%	30–50%
b -jet purity fit	3–8%	20–30%
Luminosity	3.4%	3.4%
Other sources	2%	2%

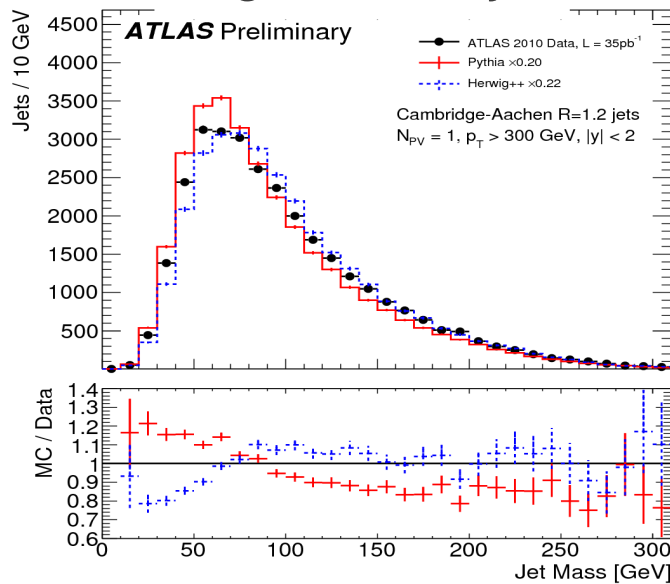
Measurement is systematics dominated

- Jet Energy Scale uncertainty
- B-tagging efficiency and purity determination

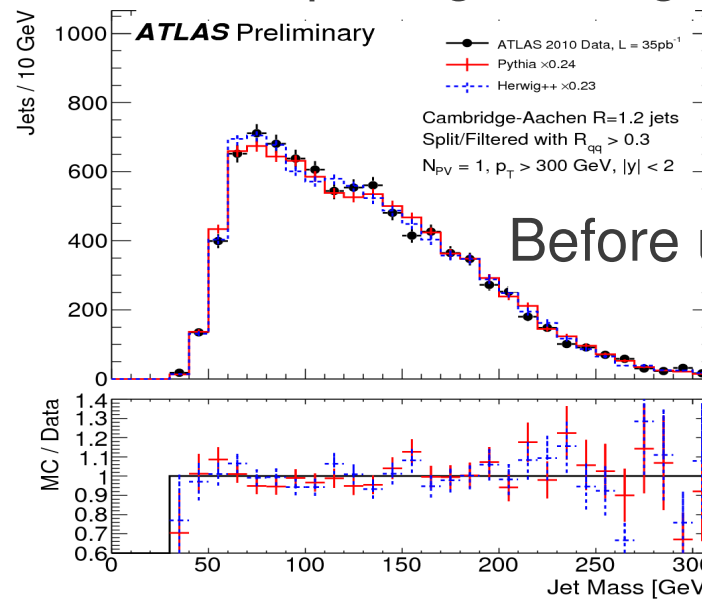
- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions

- Since we use jet mass and substructure, need to have confidence that MC describes such quantities well
- Show results for Cambridge-Aachen $R=1.2$ jets with $p_T > 300 \text{ GeV}$ and $|y| < 2$, with 2010 data
- Dominant systematics: Jet Energy and Mass Scale, Jet Energy and Mass Resolution (non-standard algorithm, default calibrations don't have to work)
 - JES, JMS: in-situ validation using track-jets, since cannot use balance techniques for mass $\rightarrow 5\text{-}7\%$
 - JER, JMR: use MC $\rightarrow 20\text{-}30\%$
- Bin-by-bin unfolding

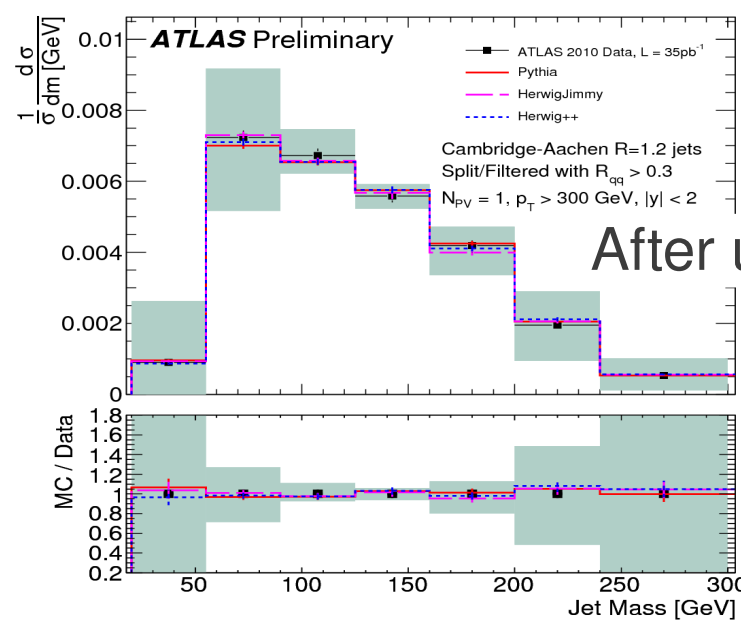
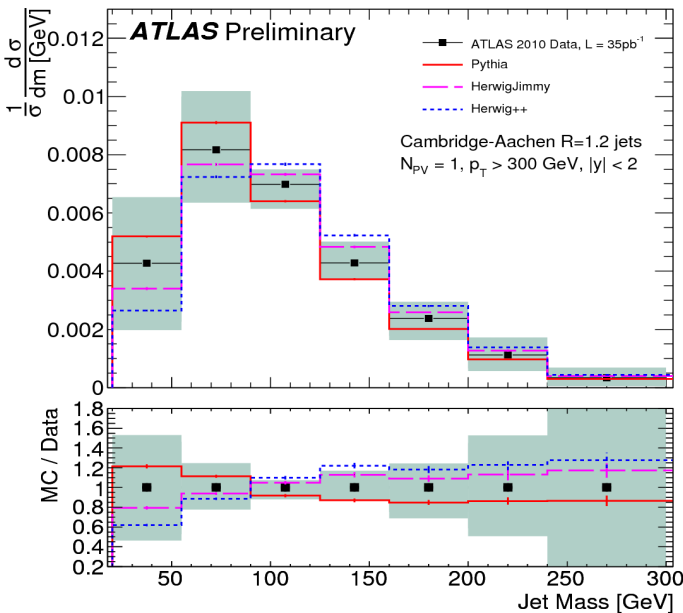
Original C-A jet



After splitting/filtering

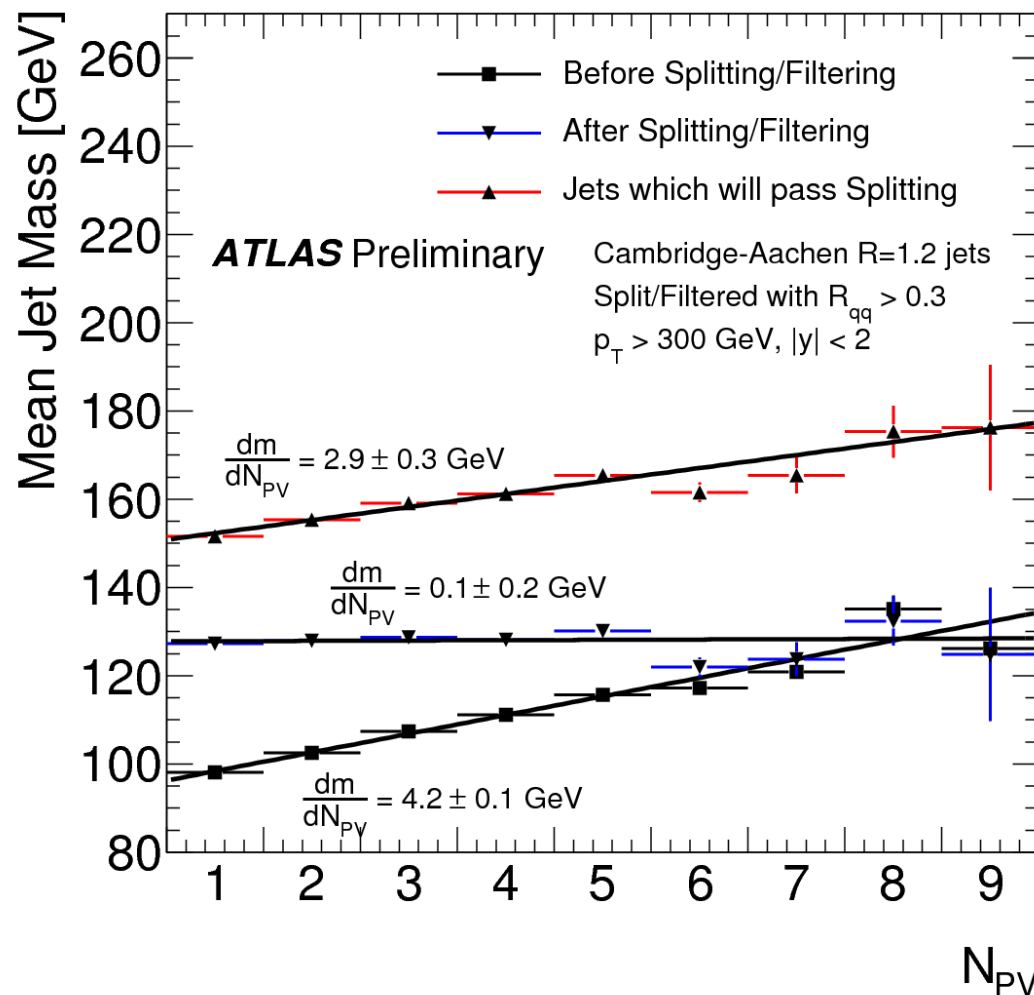


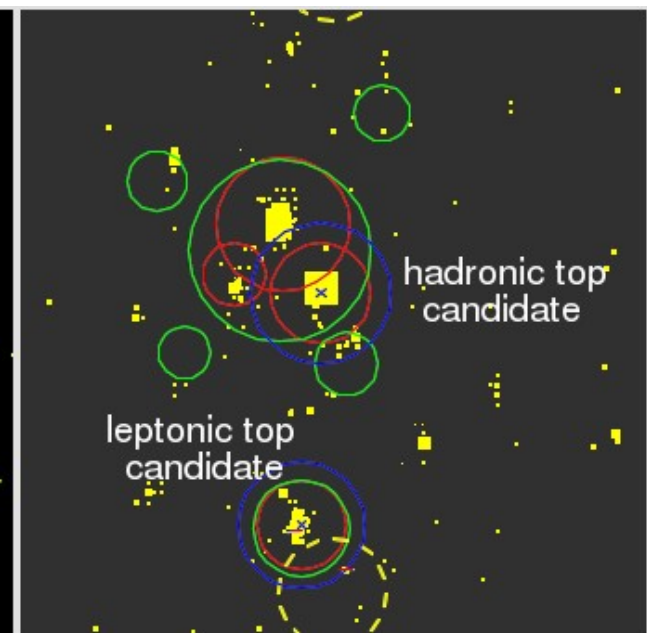
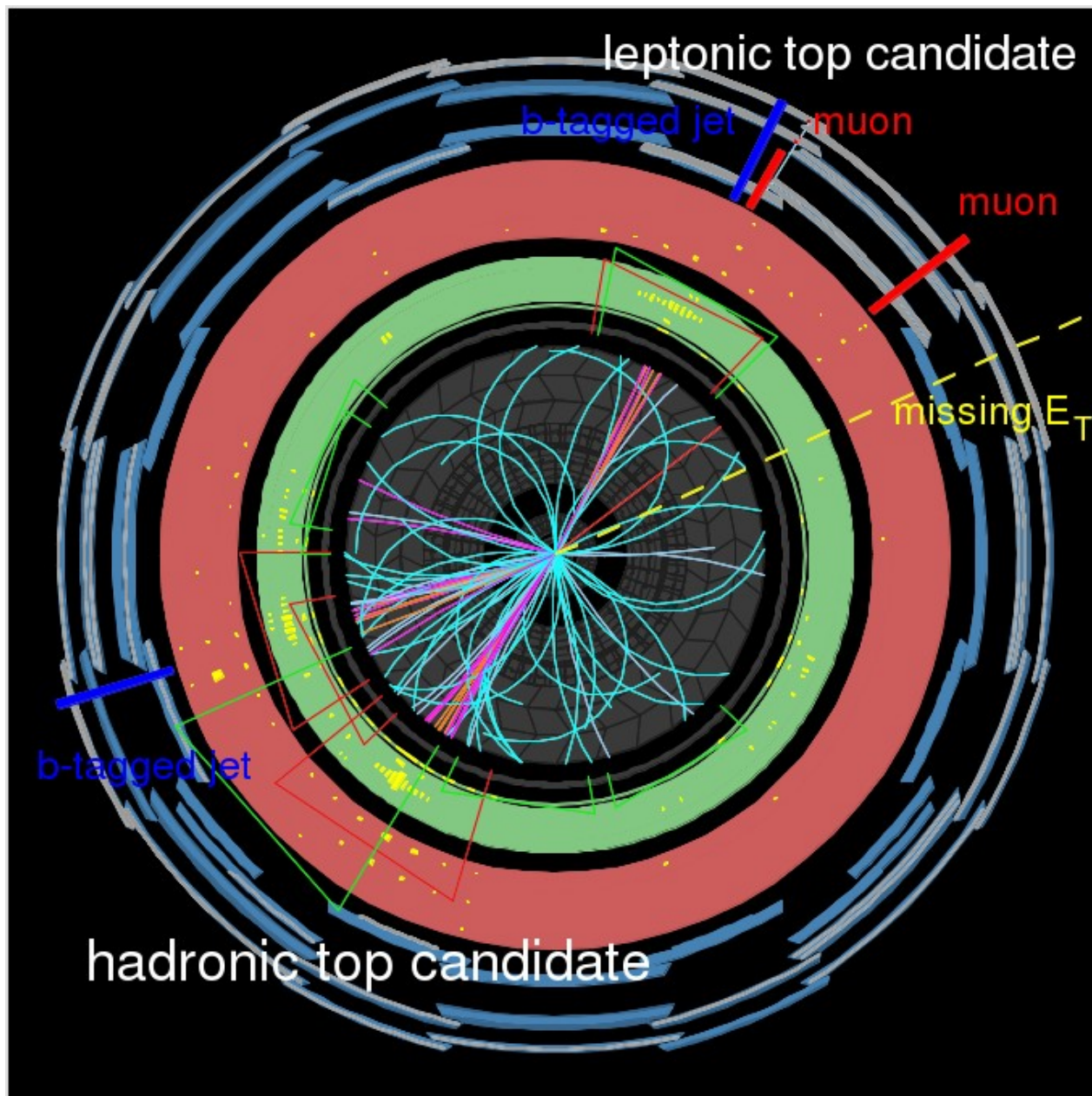
Before unfolding (detector level)



After unfolding (particle level)

- Other pp interactions in the event (== pileup) can bias the jet energy measurement, uncertainty and substructure (soft products included in reconstructed jet)
- Previous distributions without pileup (require nVertices=1). But filtering reduces dependence on pileup drastically, at least for 2010 levels.





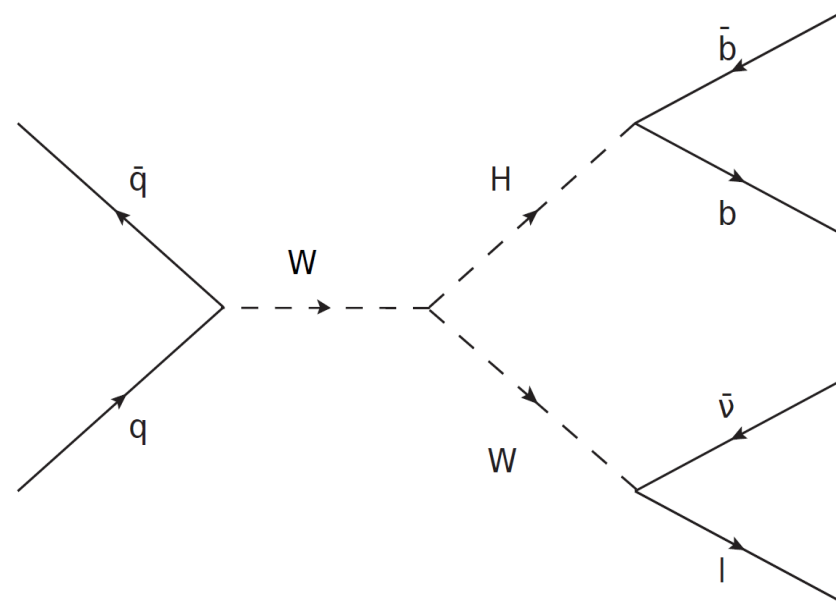
Run Number: 167576, Event Number: 106929590

Date: 2010-10-24 12:10:09 EDT

Red – antiKt04, blue – antiKt06, green – antiKt1

- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions

- Summer 2011 results shown (1.04 fb^{-1})
- $\sigma(WH) \sim 2 \times \sigma(ZH)$
- Use high- p_T leptons to trigger and identify candidate events ($W \rightarrow l\nu$, $Z \rightarrow ll$, $l=e,\mu$)
- Look for two additional b-tagged jets in the event and reconstruct the Higgs mass as m_{bb}
- Backgrounds: Vbb (irreducible), $t\bar{t}$, diboson (irreducible), QCD



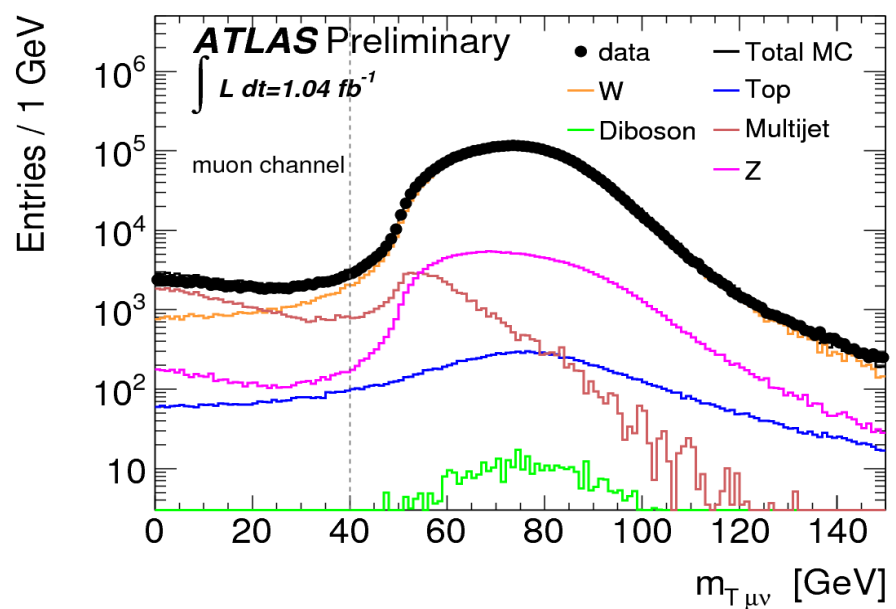
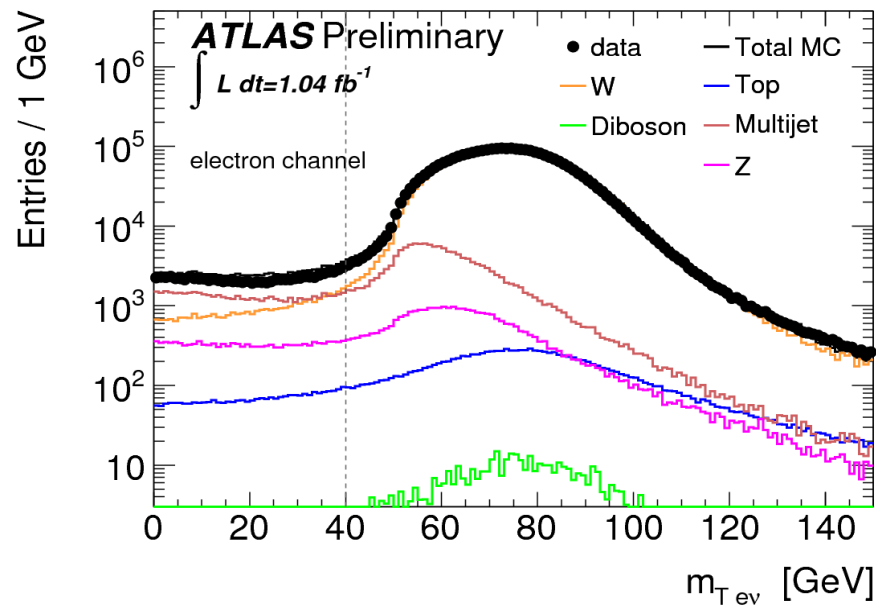
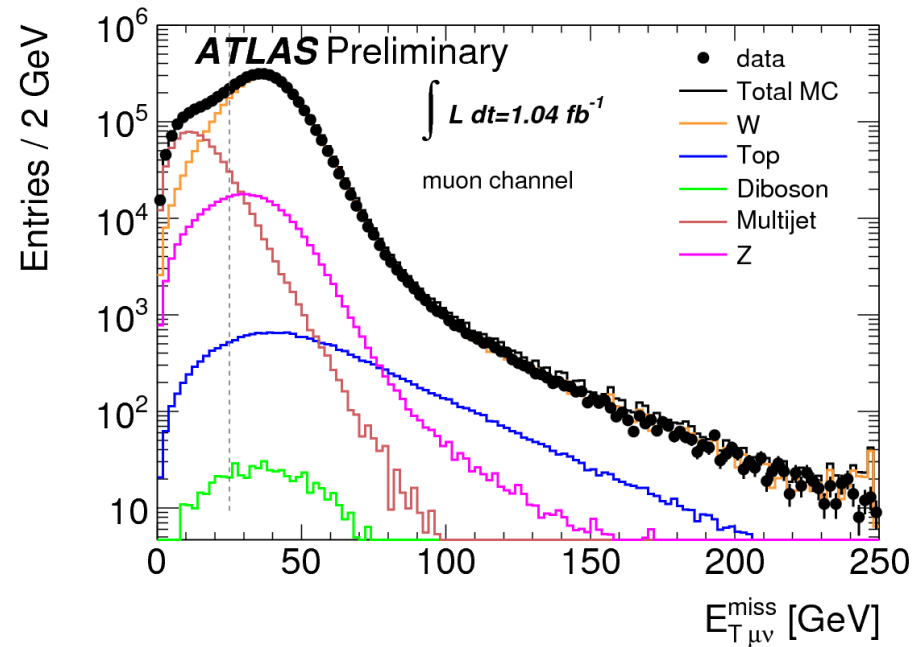
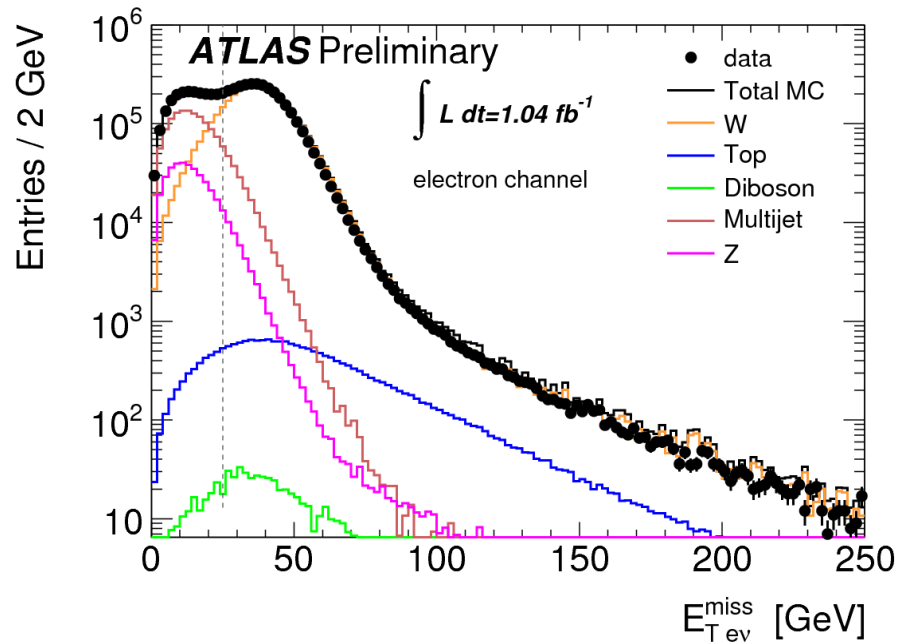
WH $\rightarrow$ bb (in general tighter)

- One good lepton in the event
- $E_{\text{T}}^{\text{miss}} > 25\text{GeV}$, $m_{\text{T}} > 40\text{GeV}$
- Exactly two jets in the event, both b-tagged – to reduce ttbar
- Good quality jets, reconstructed with the antiKt algorithm with $R=0.4$
- B-tagging: SV+IP3D combined tagger
- Single-lepton and di-electron (Z channel) triggers
- Primary vertex with at least 3 tracks

ZH $\rightarrow$ bb

- Two good leptons in the event at m_z
- $E_{\text{T}}^{\text{miss}} < 50\text{GeV}$
- At least two jets in the event, the two leading ones b-tagged

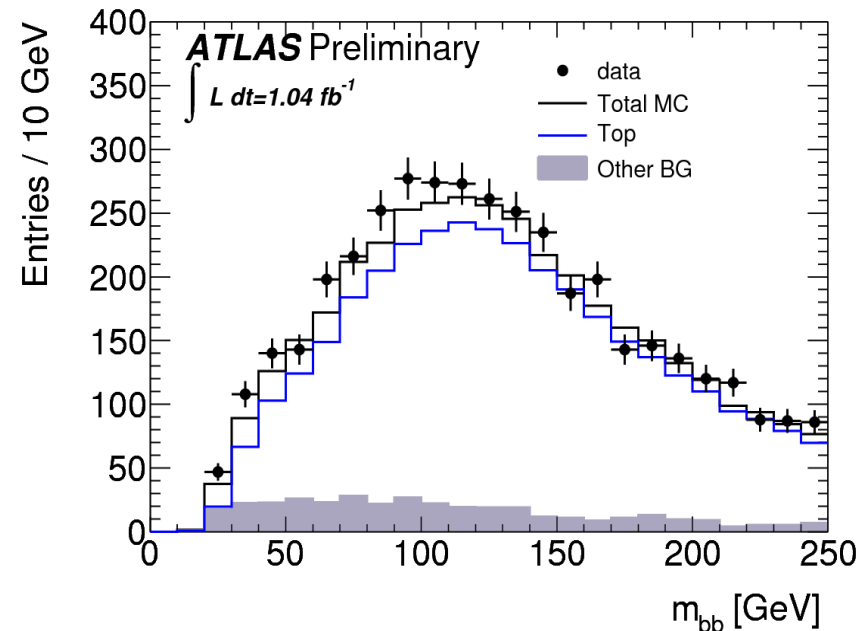
Some distributions...



ttbar and W+jets: shapes from MC (ttbar) and pre-b-tagged data (W+jets). Normalization from simultaneous fit on data using m_{bb} sidebands

- $m_{bb} < 80 \text{ GeV}$ – mainly ttbar
- $140 \text{ GeV} < m_{bb} < 250 \text{ GeV}$ – mainly W+jets

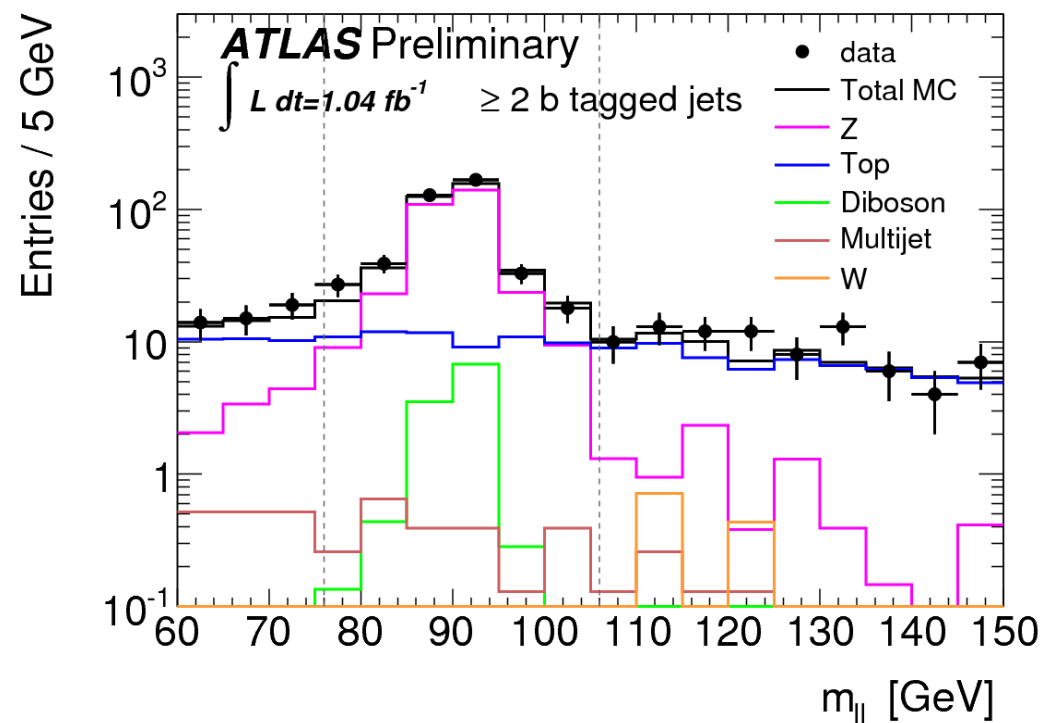
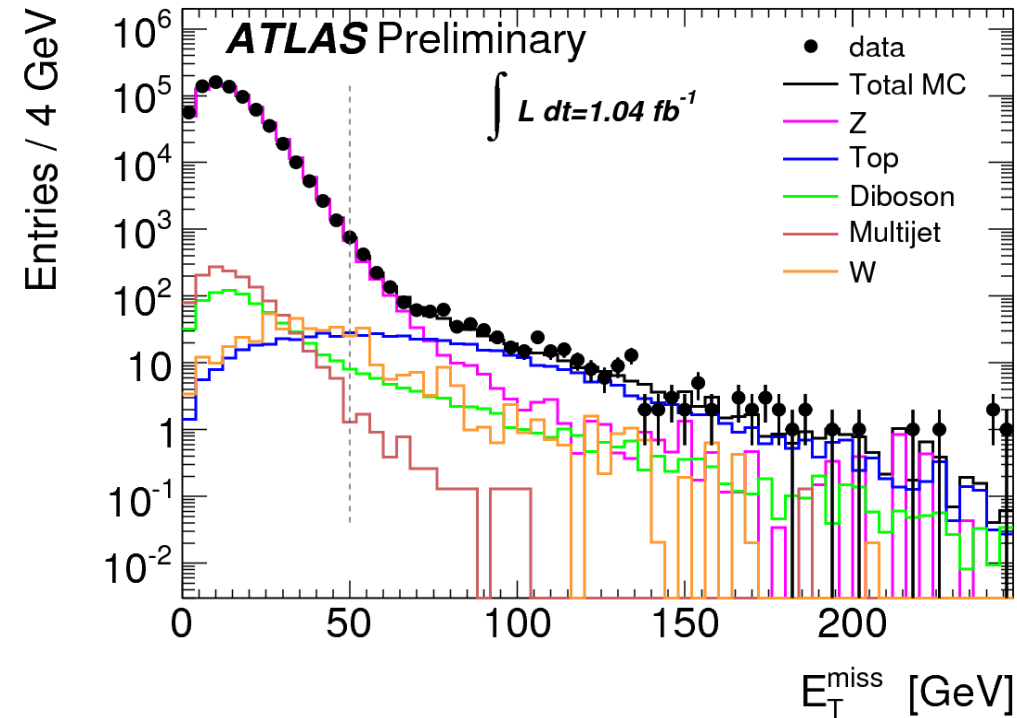
ttbar cross-check: m_{bb} for events with up to 3 jets (nJet=3 bin dominated by ttbar)



WZ and WW: irreducible background. Too small to measure from data, use MC

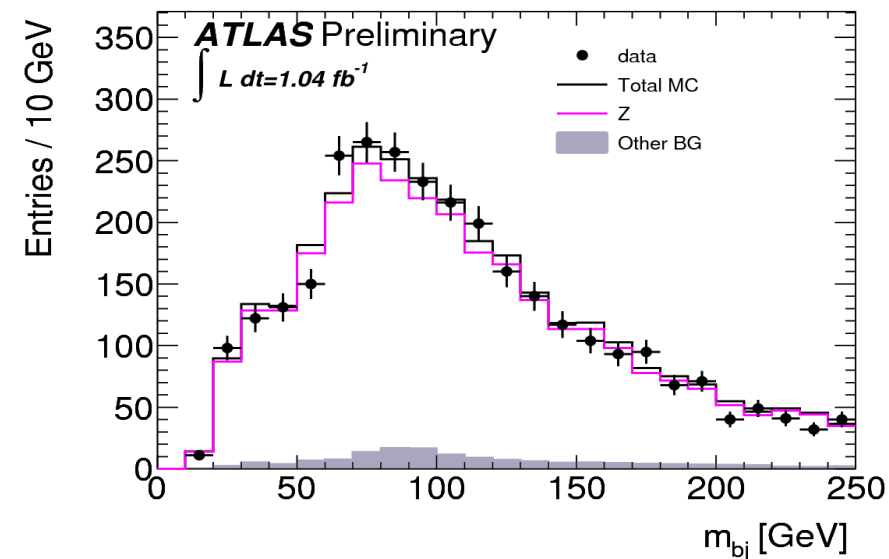
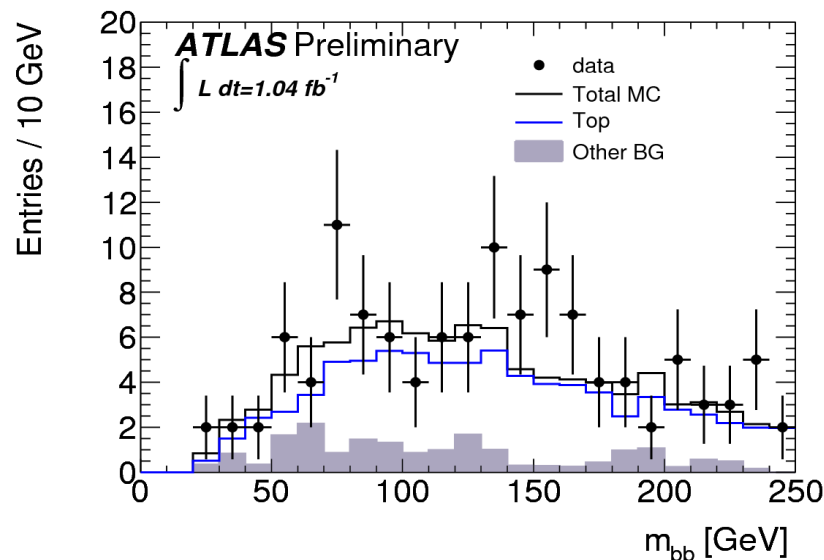
QCD multijet (fake or true leptons making a “W”): template method for e, MC for μ (negligible). Derive templates from data enriched in fake electrons, fit E_t^{miss} distribution and use normalization.

Some distributions...



Z+jets: shape from MC, normalization from data using m_{bb} sidebands ($m_{bb} < 80 \text{ GeV}$ and $140 \text{ GeV} < m_{bb} < 250 \text{ GeV}$)

- Cross-check: di-jet mass for only one b-tagged

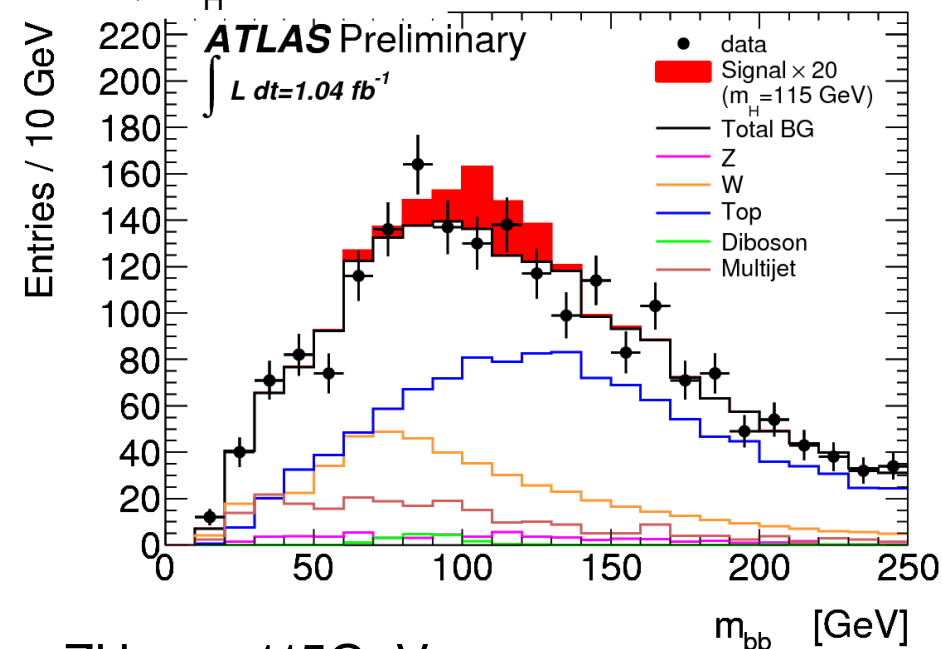


Top (t $\bar{t}$ and single): use MC, verify normalization in data using m_{ll} sidebands ($60 \text{ GeV} < m_{ll} < 76 \text{ GeV}$ and $106 \text{ GeV} < m_{ll} < 150 \text{ GeV}$)

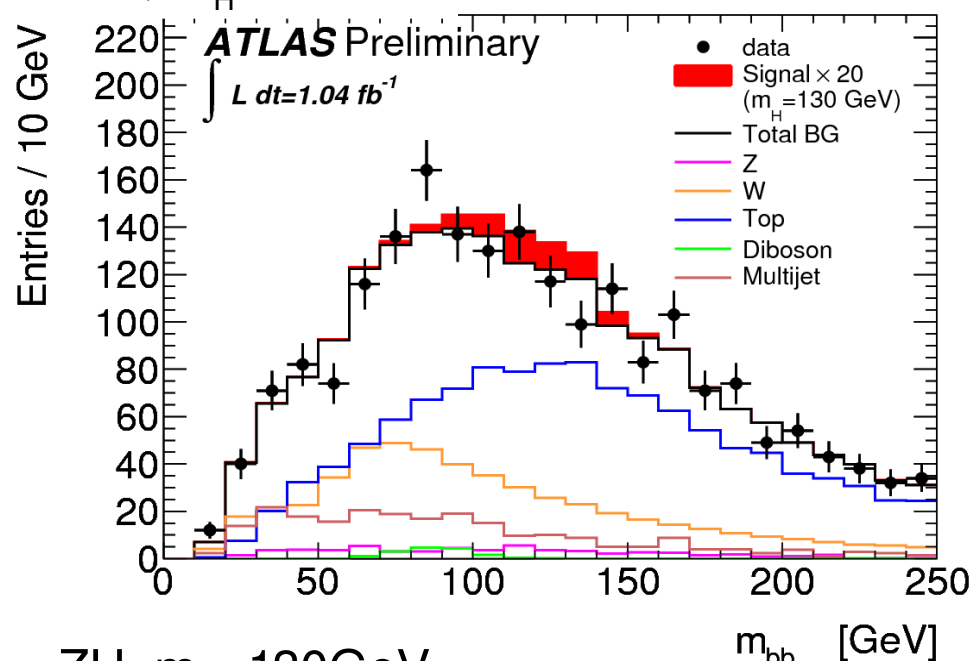
ZZ: irreducible background. Too small to measure from data, use MC

QCD multijet (fake or true leptons making a “Z”): template method for e, MC for μ (negligible). Derive templates from data enriched in fake electrons,⁴⁰ fit di-lepton mass distribution and use normalization.

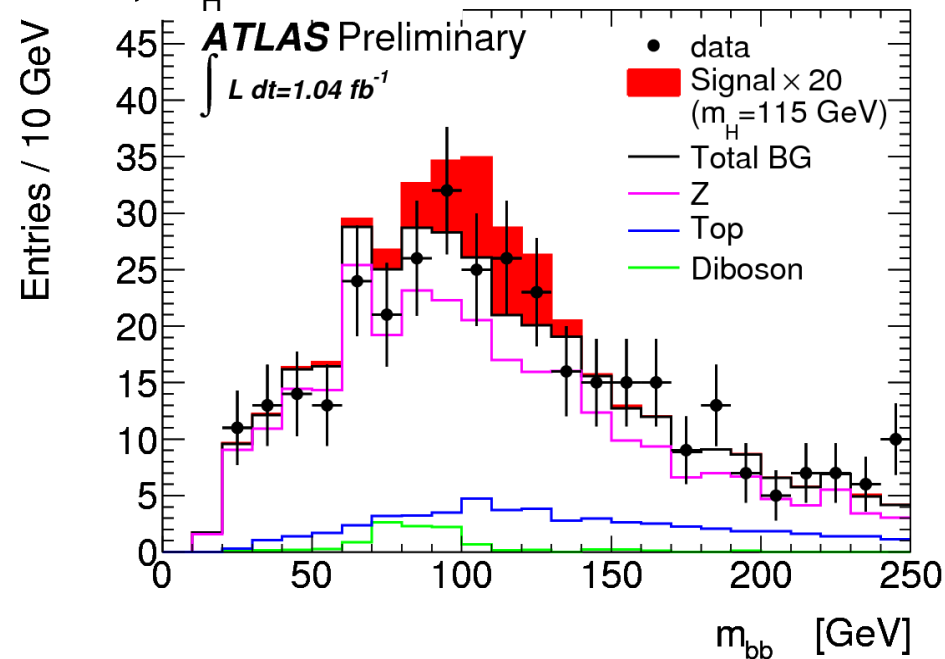
WH, $m_H=115\text{GeV}$



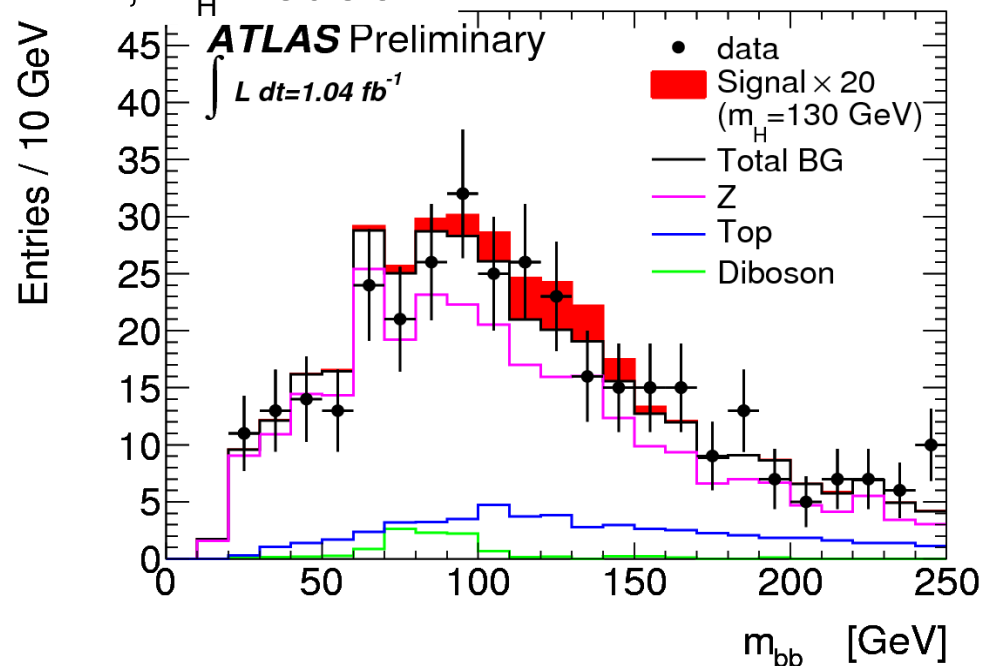
WH, $m_H=130\text{GeV}$

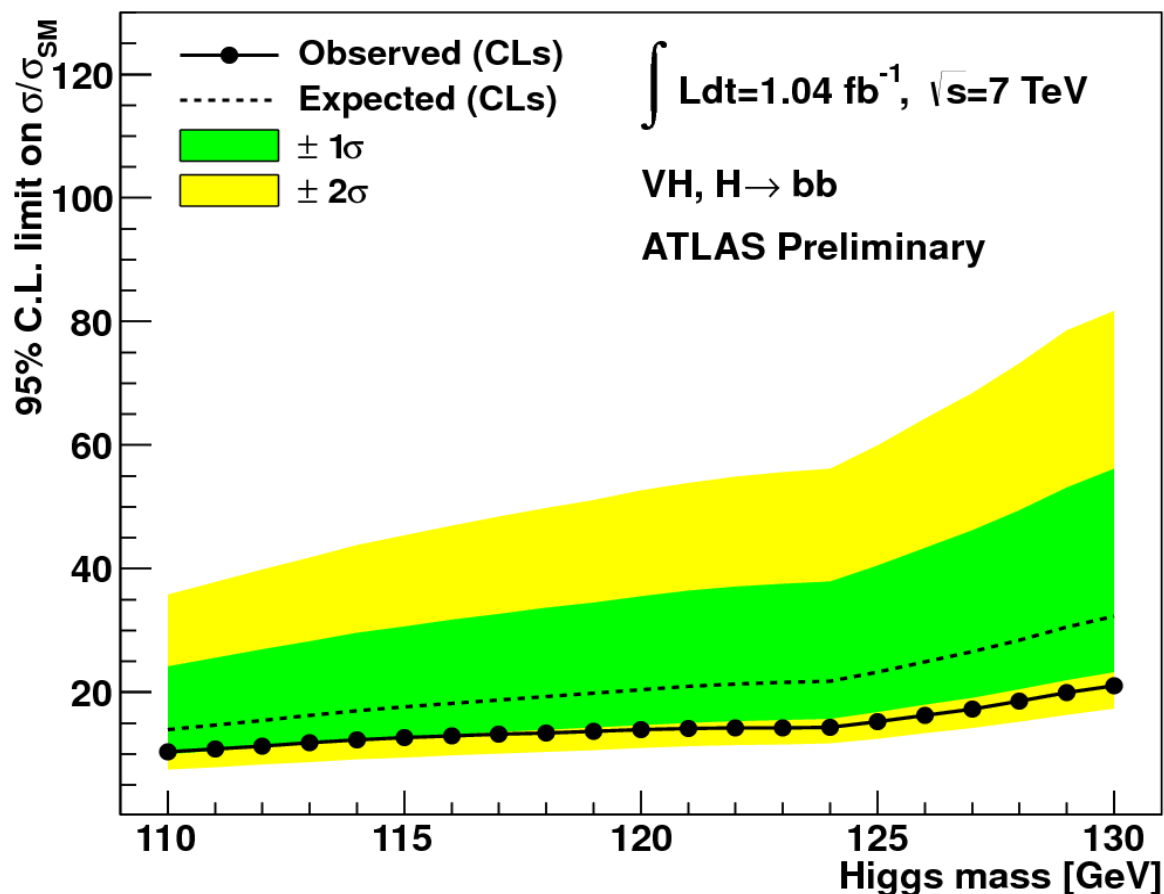


ZH, $m_H=115\text{GeV}$



ZH, $m_H=130\text{GeV}$





Expected and observed exclusion limits for the ZH and WH channels combined. Using profile-likelihood method with CL_s

→ all the observed points stay lower than the expected due to the very wide mass distribution ($\sigma_{mH} \sim 18\text{-}20\text{ GeV}$), dominated by out-of-cone losses

- Detector-related – mainly b-tagging, JES, pileup
- Higgs cross-section
- Background estimates
- Luminosity

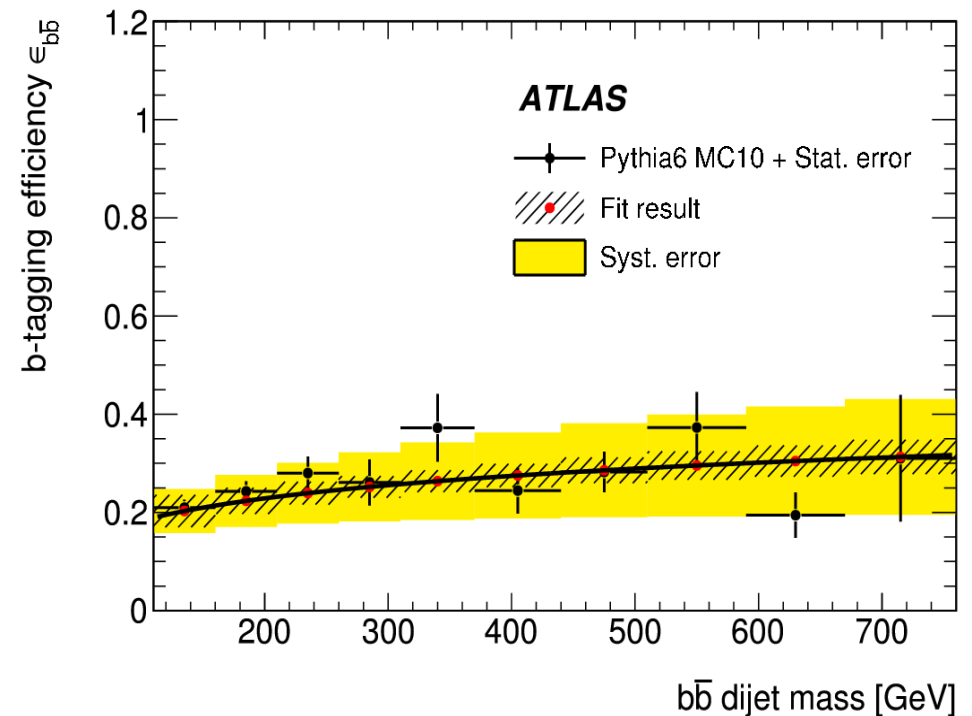
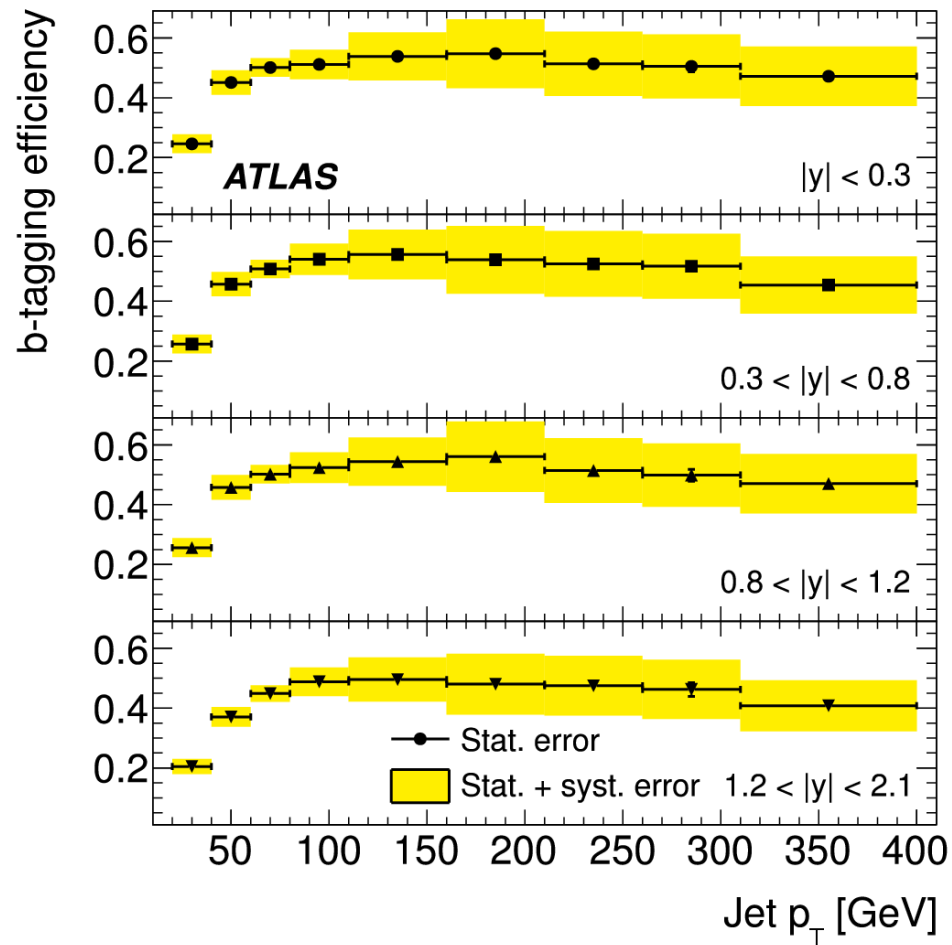
Source of Uncertainty	Treatment in analysis	
	<i>ZH</i>	<i>WH</i>
Luminosity	3.7%	3.7%
Higgs boson cross-section	5%	5%
Background norm. and shape:		
Top	9%	6%
Z+jets	9% plus shape	9%
W+jets	negligible	14% plus shapes
ZZ	11%	negligible
WZ	11%	11%
WW	negligible	11%
QCD multijets	100%	50%

Source of Uncertainty	Effect on $ZH \rightarrow \ell\ell b\bar{b}$ signal		Effect on $WH \rightarrow \ell\nu b\bar{b}$ signal	
	$m_H = 115 \text{ GeV}$	$m_H = 130 \text{ GeV}$	$m_H = 115 \text{ GeV}$	$m_H = 130 \text{ GeV}$
Electron Energy Scale	< 1%	< 1%	1%	1%
Electron Energy Resolution	< 1%	< 1%	1%	1%
Muon Momentum Resolution	1%	3%	4%	1%
→ Jet Energy	9%	7%	1%	3%
Jet Energy Resolution	< 1%	< 1%	1%	1%
Missing Transverse Energy	2%	2%	2%	3%
→ <i>b</i> -tagging Efficiency	16%	17%	16%	17%
<i>b</i> -tagging Mis-tag Fraction	< 1%	< 1%	3%	3%
Electron Efficiency	1%	1%	1%	1%
Muon Efficiency	1%	1%	1%	1%
Luminosity	4%	4%	4%	4%
Higgs Cross-section	5%	5%	5%	5%

- Motivation – why boosted $VH \rightarrow bb$?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted $VH \rightarrow bb$ search
 - Inclusive b-jet and bb cross section
 - Jet substructure measurements
 - Un-boosted (inclusive) $VH \rightarrow bb$
- Conclusions

- Boosted $VH \rightarrow bb$ can contribute valuable sensitivity to low-mass Higgs searches
- ATLAS has proven good understanding of jets, b-tagging and even jet substructure, as shown by early SM and performance measurements
- Inclusive $VH \rightarrow bb$ measurement available, working on further improvements
- Moving on to boosted $VH \rightarrow bb$:
 - Can probably relax some cuts, but...
 - Estimating low backgrounds with low data statistics challenging
 - Special jet energy and mass systematics required
 - B-tagging in busy boosted jet environment
 - Use boosted $Z \rightarrow bb$ to constrain some of those systematics

Measured in MC, then scaled to data-driven efficiency measurement with official b-tagging group Scale Factors



WH $\rightarrow$ bb

- Electrons: “tight” quality, $p_T > 25\text{GeV}$, $|\eta| < 2.47$, isolation, $|d_o| < 0.1\text{mm}$, $|z_o| < 10\text{mm}$
- Muons: $p_T > 25\text{GeV}$, $|\eta| < 2.4$, isolation, $|d_o| < 0.1\text{mm}$, $|z_o| < 10\text{mm}$, $\Delta R > 0.4$ from jets
- One lepton in the event
- $E_T^{\text{miss}} > 25\text{GeV}$, $m_T > 40\text{GeV}$
- Exactly two jets in the event, both b-tagged
- Jets (antiKt algorithm with $\Delta R = 0.4$): jet cleaning, $p_T > 25\text{GeV}$, Jet Vertex Fraction > 0.75 , $|\eta| < 2.5$, $\Delta R > 0.4$ from electrons
- B-tagging: SV+IP3D combined tagger, at 70% efficiency – 50 light rejection working point on $t\bar{t}$ simulation
- Single-lepton and di-electron (Z channel) triggers
- Primary vertex with at least 3 tracks

ZH $\rightarrow$ bb

- Electrons: “medium” quality, $p_T > 20\text{GeV}$, $|\eta| < 2.47$, isolation
- Muons: $p_T > 20\text{GeV}$, $|\eta| < 2.5$, isolation
- Two same-flavor leptons in the event, with $76\text{GeV} < m_{ll} < 106\text{GeV}$
- $E_T^{\text{miss}} < 50\text{GeV}$
- At least two jets in the event, the two leading ones b-tagged