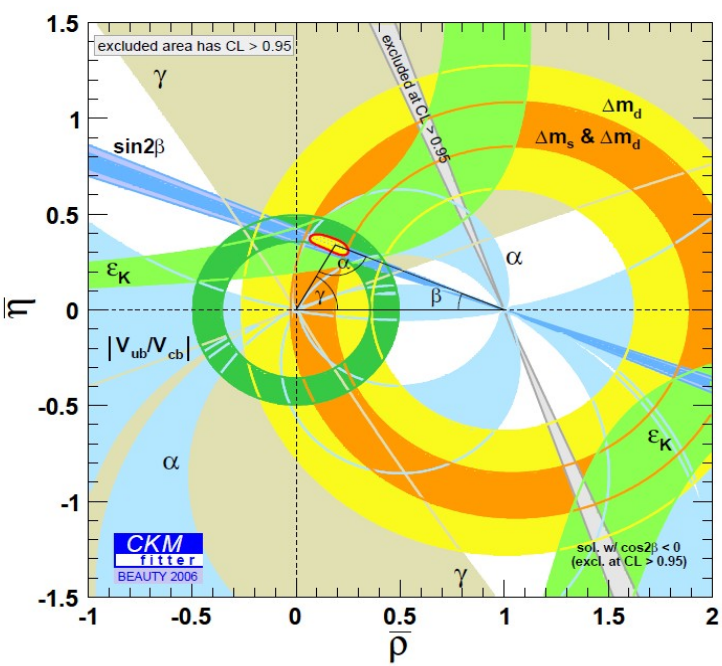
Towards a boosted VH→bb measurement with ATLAS

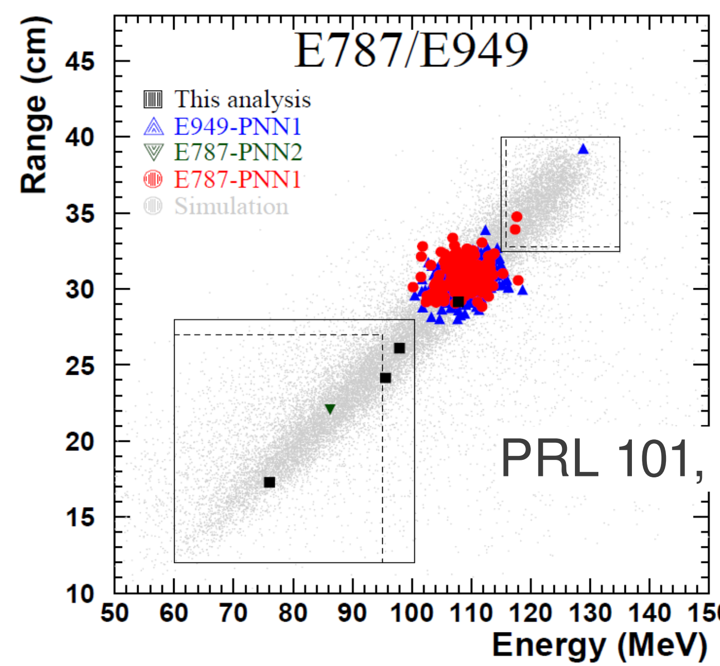
CPPM-Marseille seminar
26 January 2012

Ilektra Christidi
(CNRS CR1/CR2 candidate)

A bit about me...

- 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 $p_{\pi+}$<199MeV/c”, with E949 experiment, BNL.



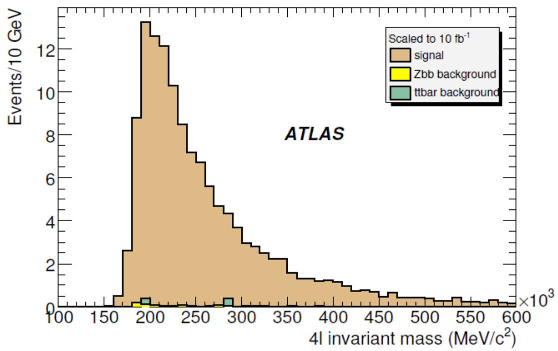


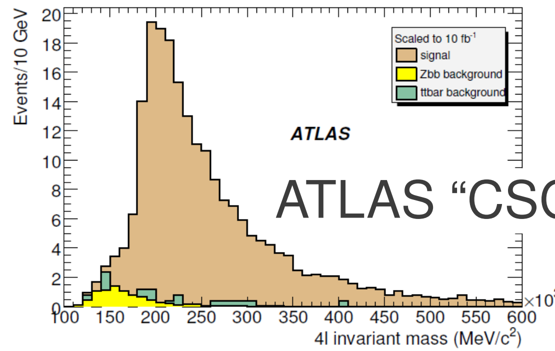
PRL 101, 191802 (2008)

A bit about me...

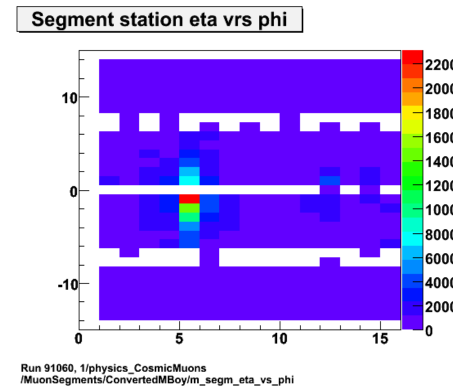
2007-2009: Marie Curie Experienced Researcher, Aristotle University of Thessaloniki, Greece, with ATLAS.

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



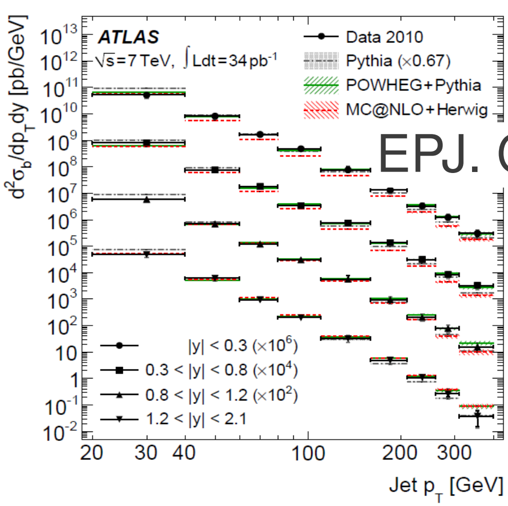


ATLAS "CSC book"

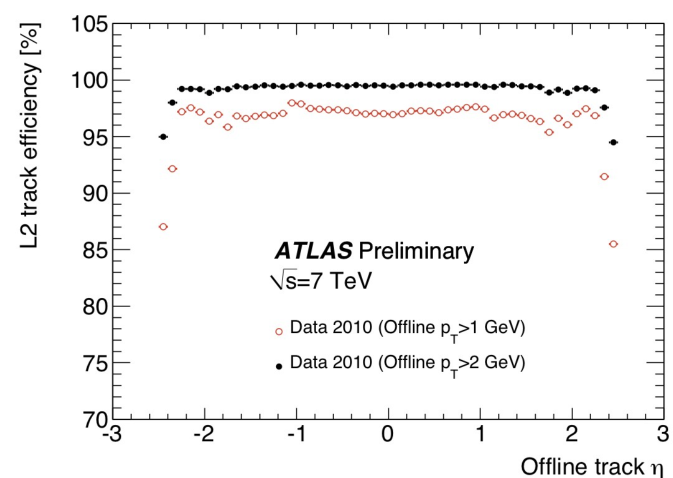


2009-2012: postdoctoral researcher partly funded by Marie Curie, UCL, UK, with ATLAS.

- 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)

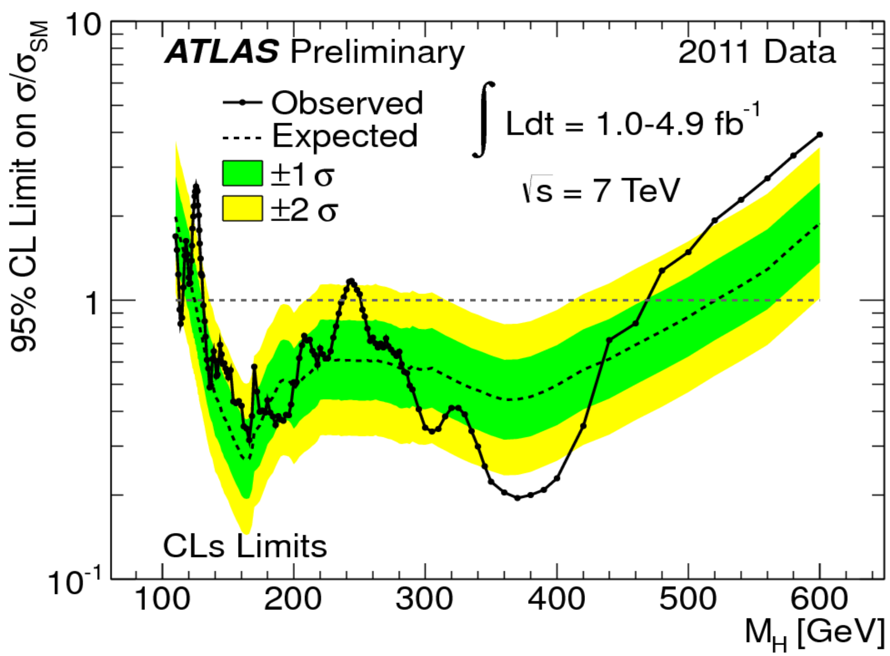


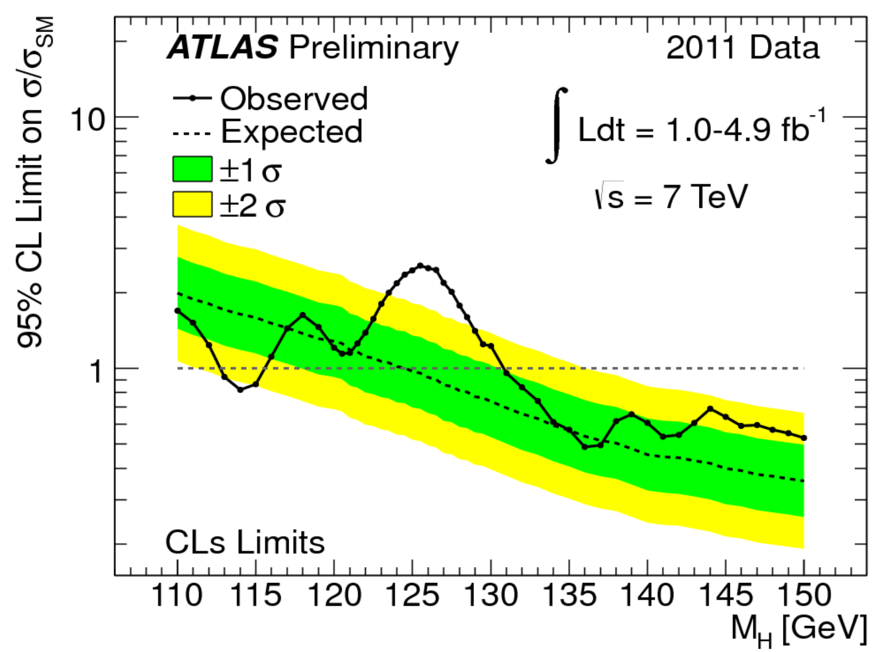
Overview

- Motivation – why boosted VH→bb ?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted VH→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→bb
    - many common elements (ATLAS-CONF-2011-103)
- Conclusions

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

Why H→bb ?





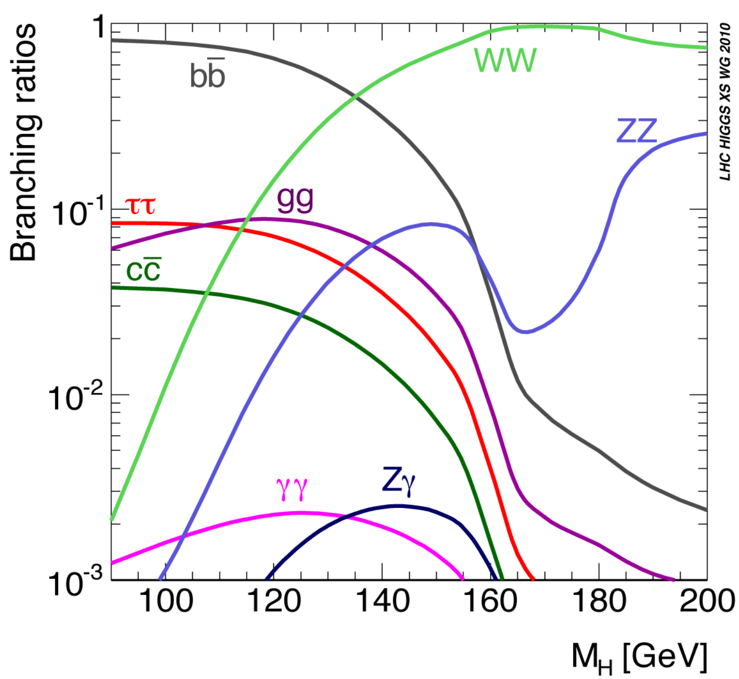
Current allowed regions (ATLAS only):

115.5 GeV < m_H < 131 GeV

237 GeV < m_H < 251 GeV

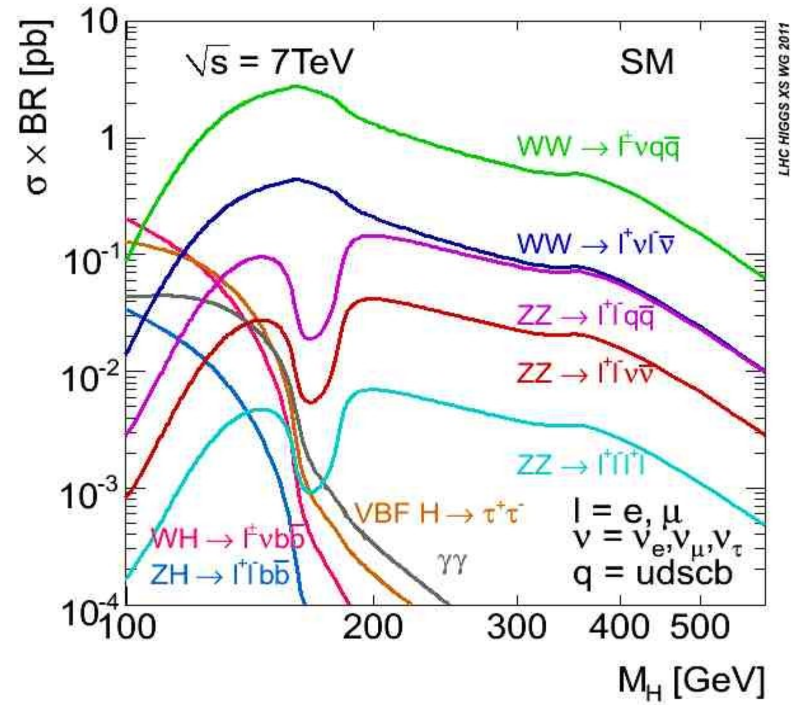
m_H > 468 GeV

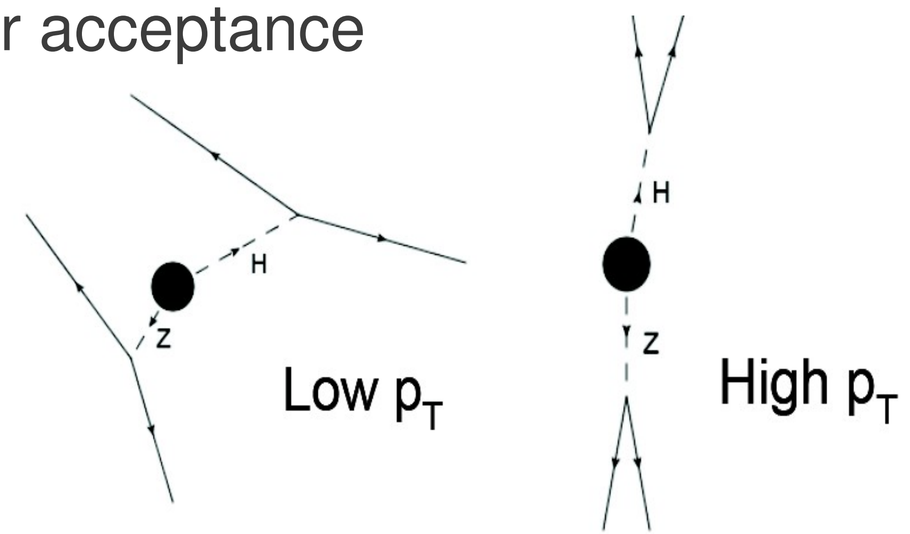
In the low-mass region, H→bb is the dominant decay channel



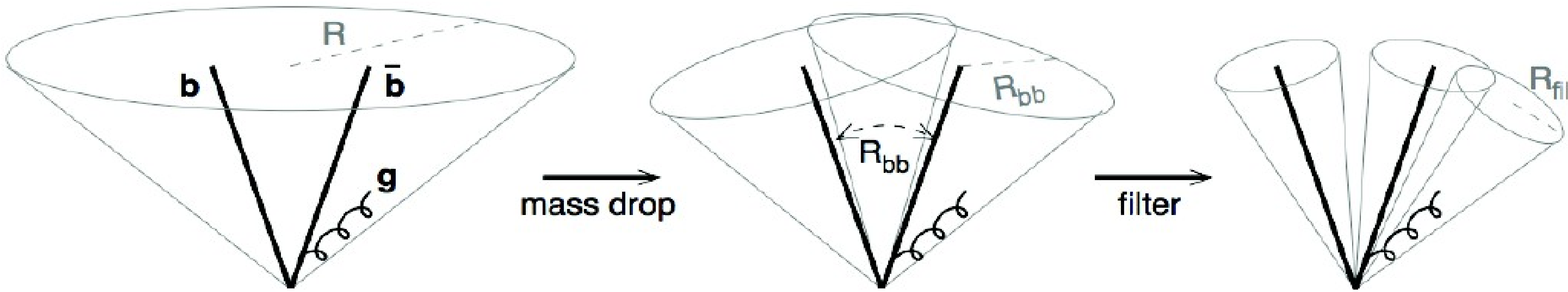
Why boosted VH→bb ?

- gg→H→bb overwhelmed by backgrounds, therefore VH (V=W or Z) associated production more sensitive
- Still, VH→bb quite background-prone.
- New idea: use only highly boosted H/V, which costs ~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





Reconstructing the boosted Higgs



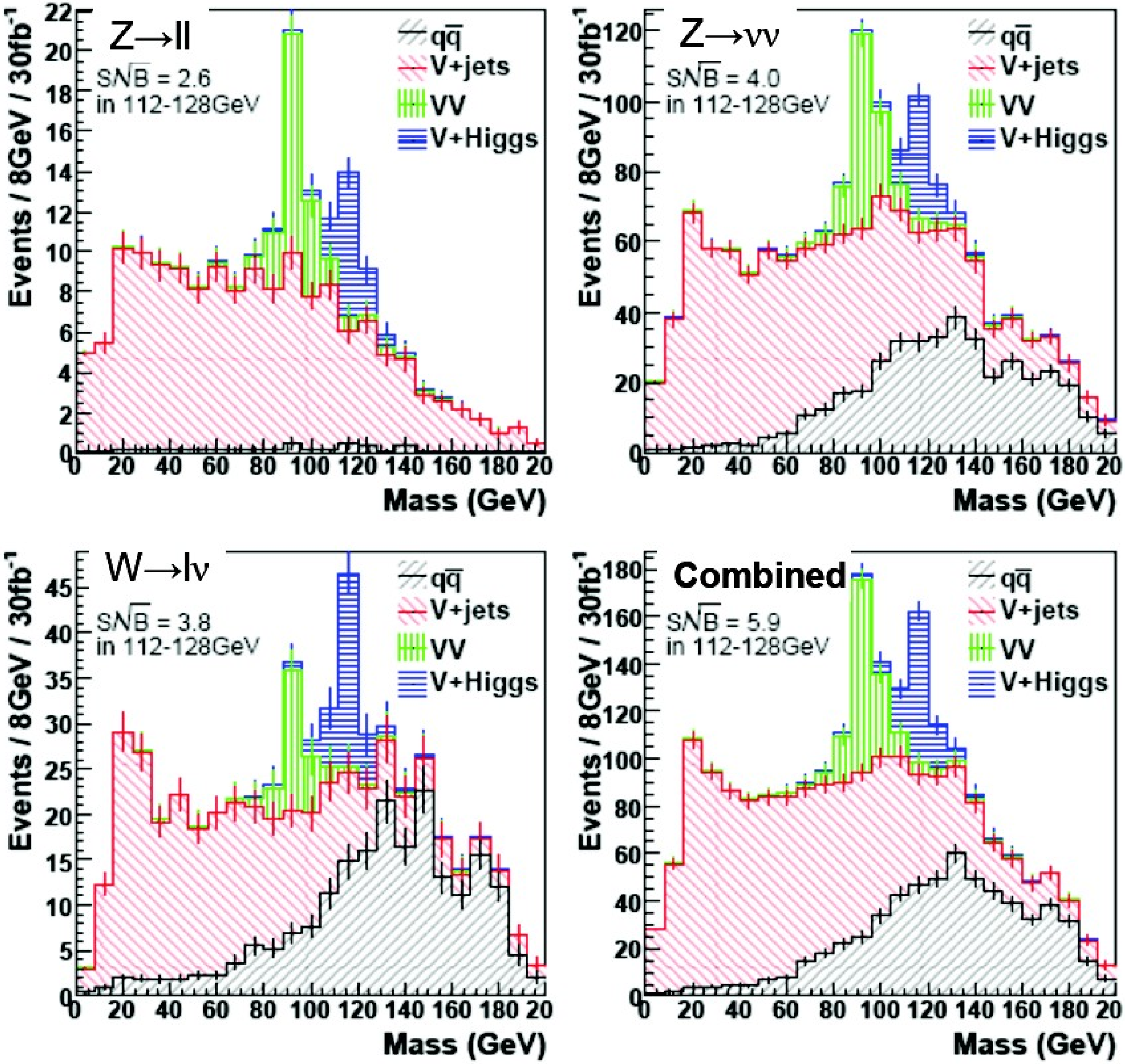
Use the angle-ordered Cambridge-Aachen (C-A) jet algorithm, which keeps track of clustering steps.

- Clustering can be undone one step at a time with an angular distance $\Delta R=\sqrt{(\Delta y^2+\Delta\varphi^2)}$ 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

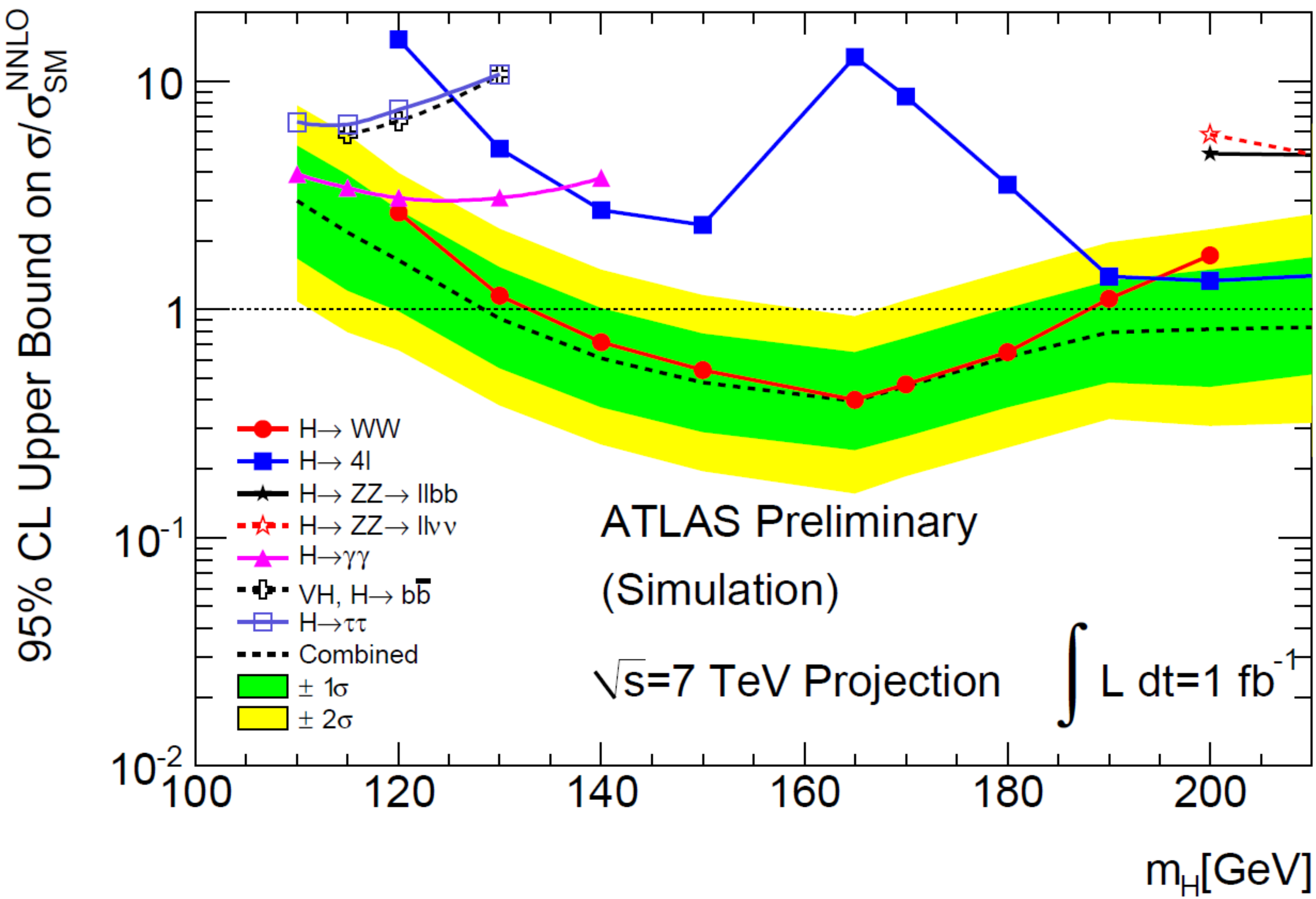
Expected performance

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



Expected performance

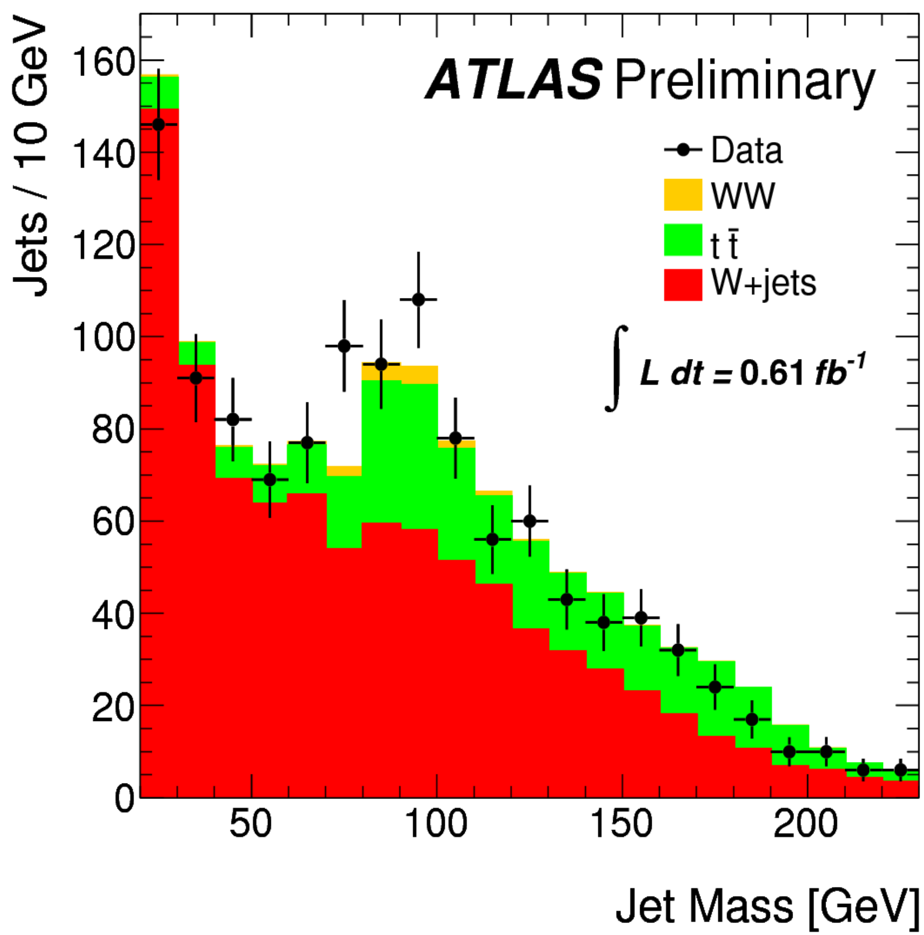
From ATLAS-PUB-2010-015 (full ATLAS detector simulation), 1 fb^{-1}



Boosted lvqq in data

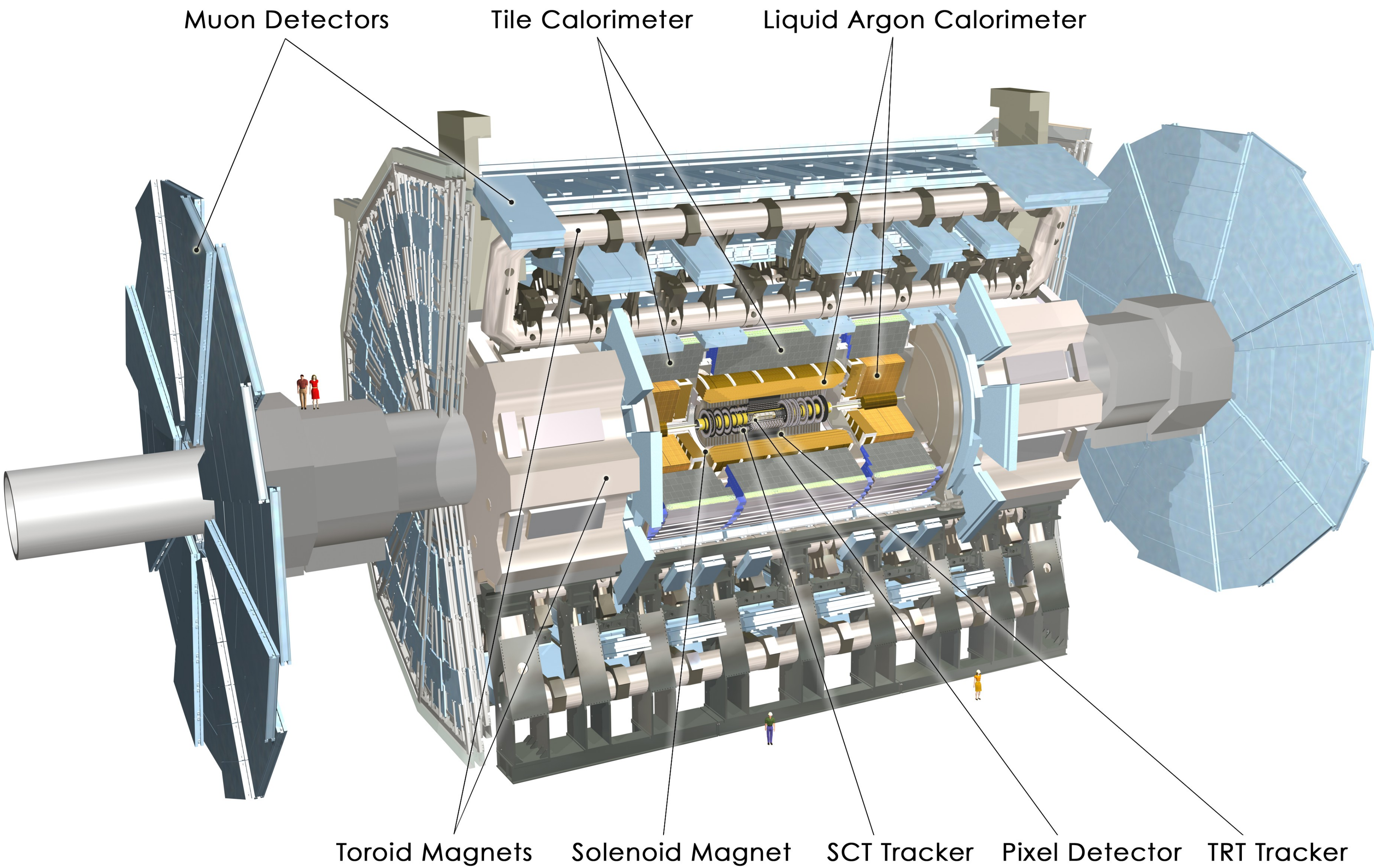
Very encouraging first result from summer 2011:

Split and filtered C-A 1.2 jet mass, in association with a W→lν with p_T>200GeV. No b-tagging applied.



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

The ATLAS detector



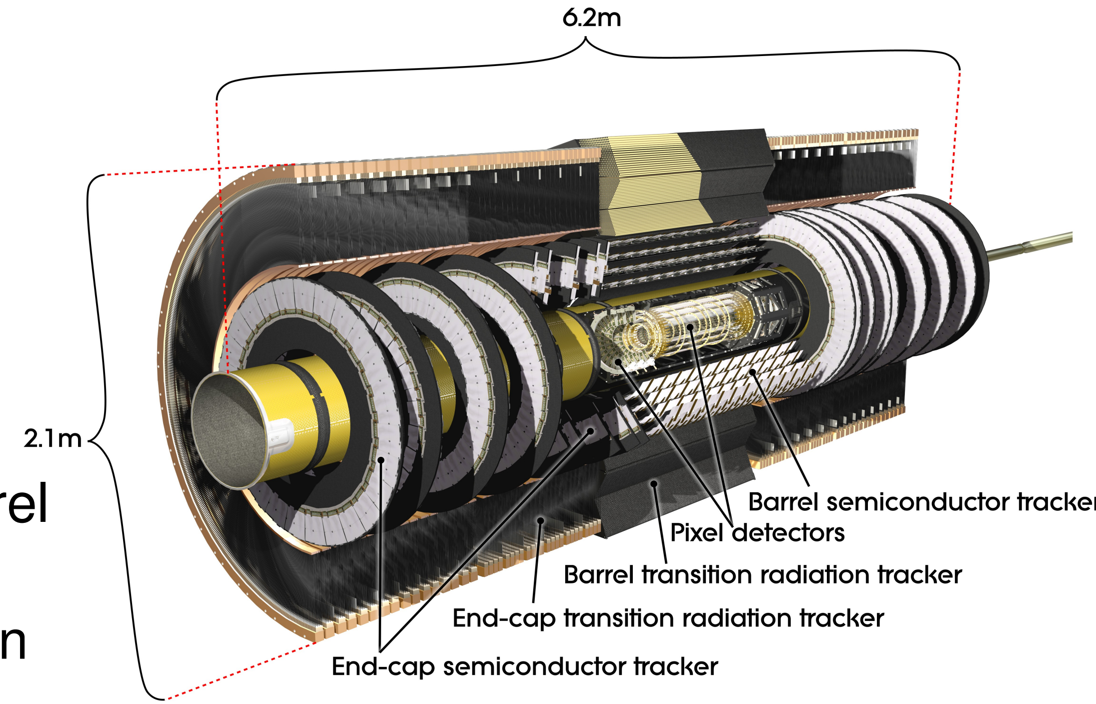
The ATLAS Inner Detector (ID)

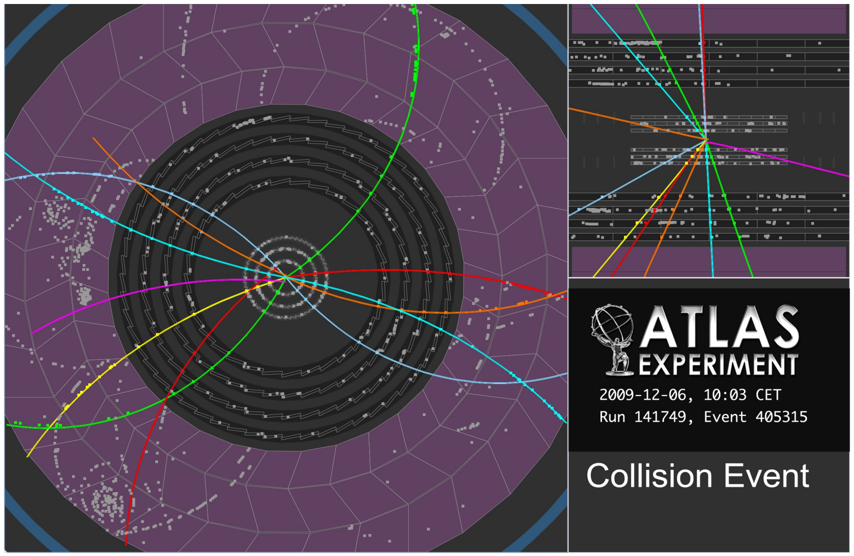
Detectors of 3 different technologies for optimal position detection and track reconstruction:

- 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.

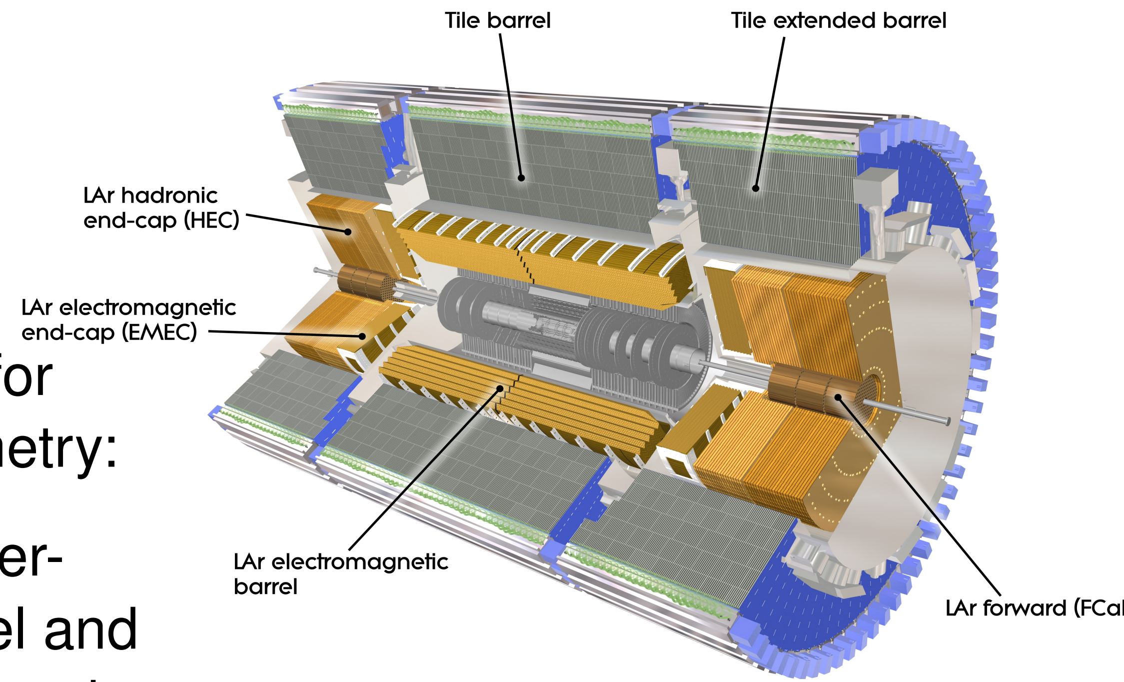
→ Track-based b-tagging happens here





http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html

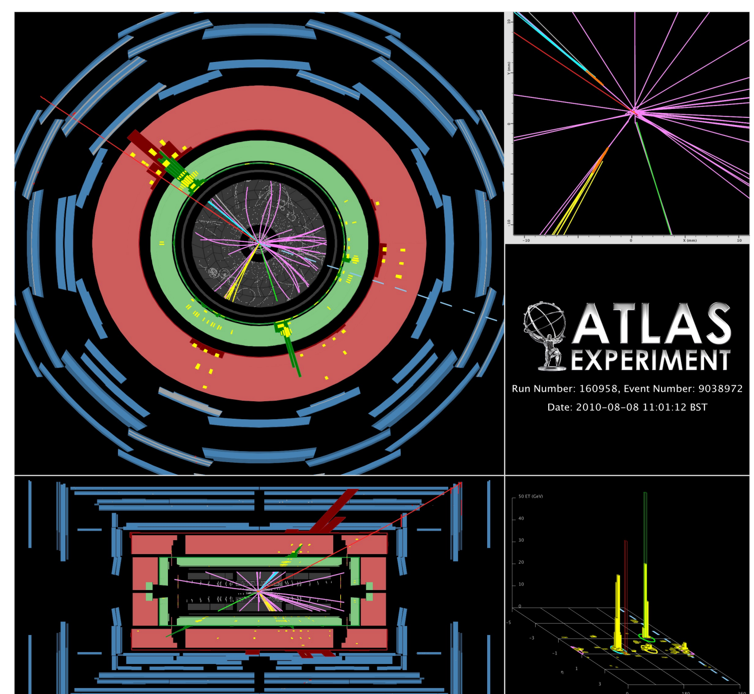
The ATLAS Calorimeters (EM and HAD)



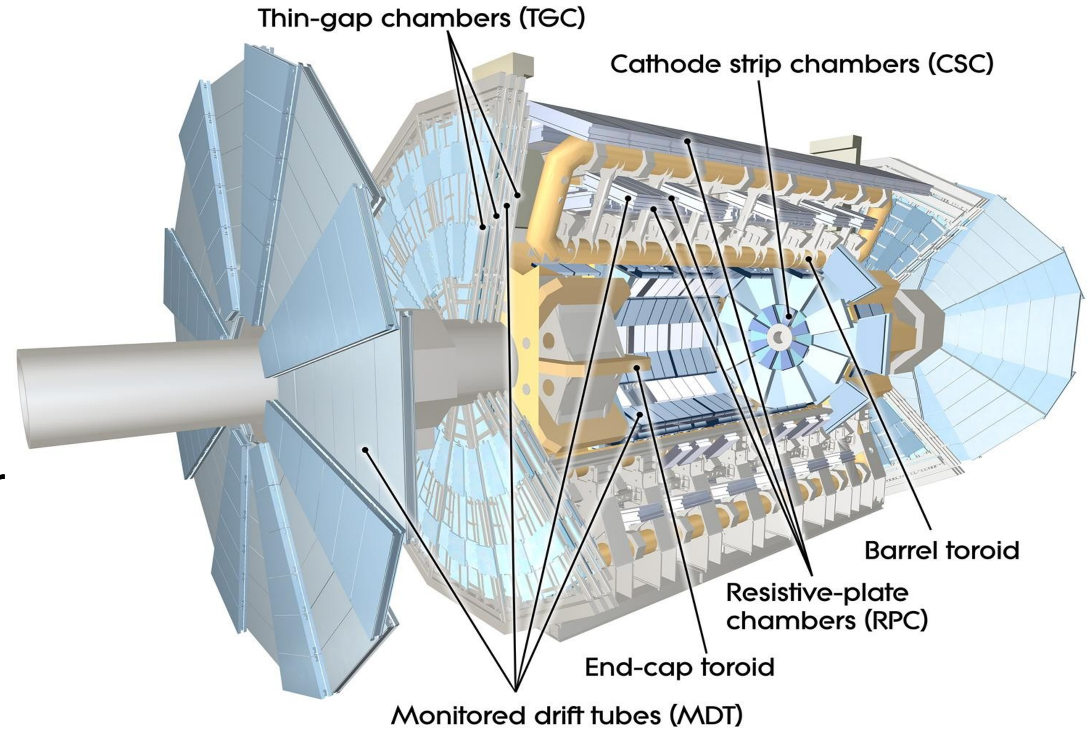
Detectors of 2 different technologies for electromagnetic and hadronic calorimetry:

- 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



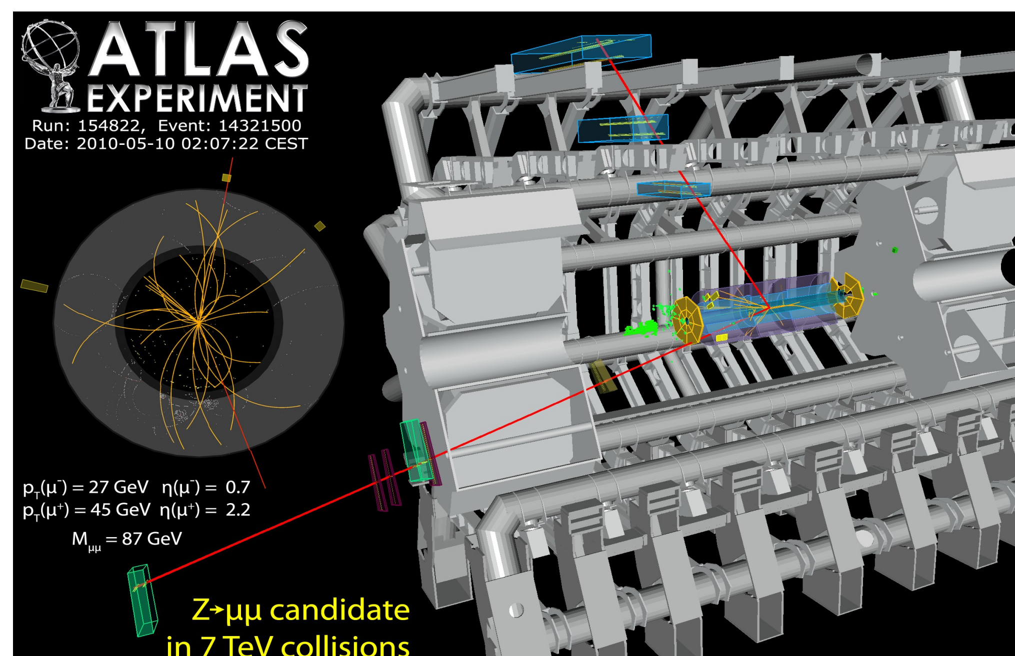
The ATLAS Muon Spectrometer (MS)



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

- Monitored Drift Tubes (**MDT**s) for precision position measurement
- Resistive Plate Chambers (**RPC**s) for triggering in the Barrel
- Thin Gap Chambers (**TGC**s) for triggering in the ECs
- Cathode Strip Chambers (**CSC**s) for tracking in the busier, high-η region

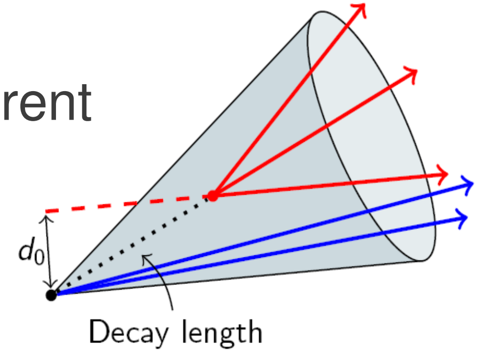
The whole structure is immersed in a 0.5T toroid magnetic field.



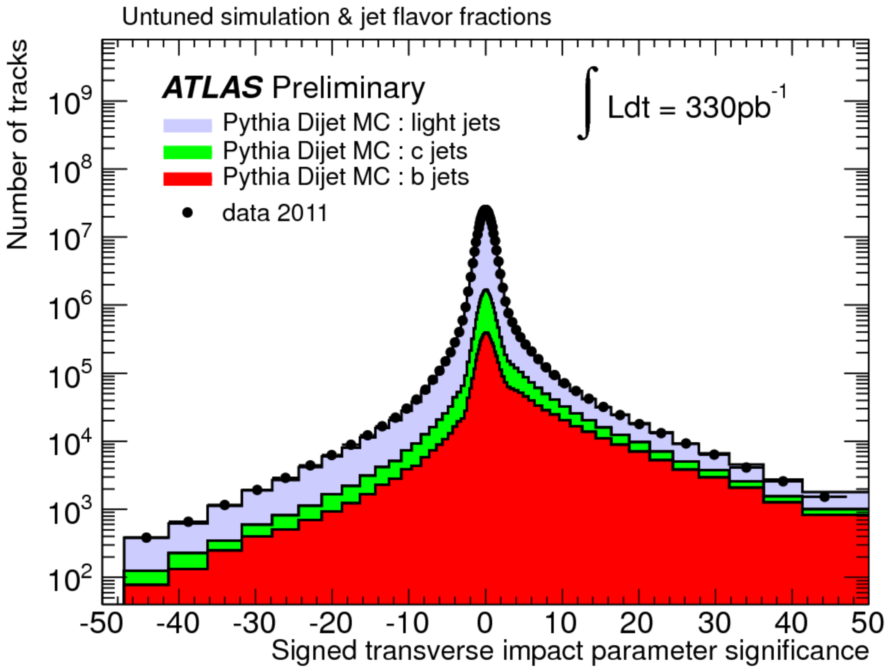
→ Muons are reconstructed in the MS and ID

Track-based b-tagging

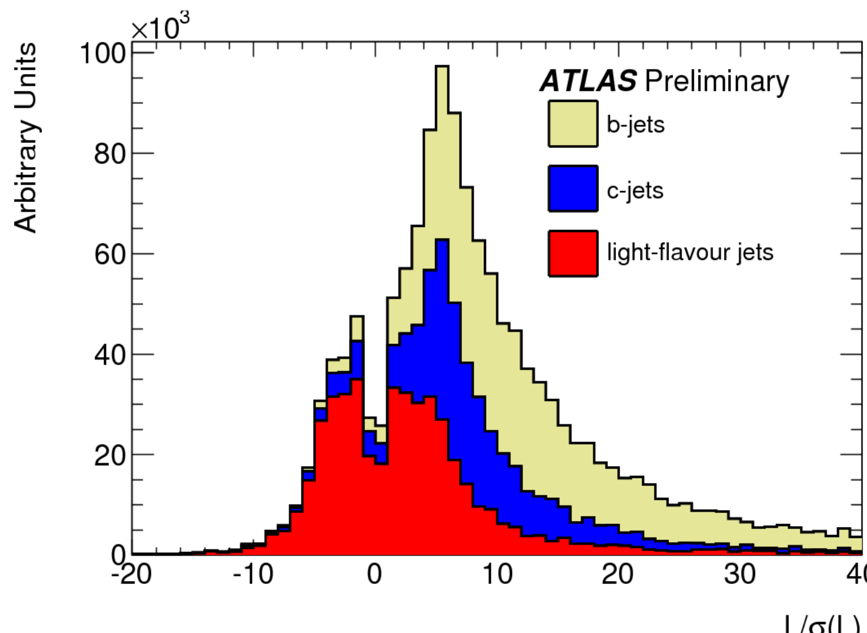
Tracks from the b-decay have different topology than prompt ones



Impact Parameter-based algorithms: exploit the fact that b-decay tracks are not compatible with Primary Vertex → 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

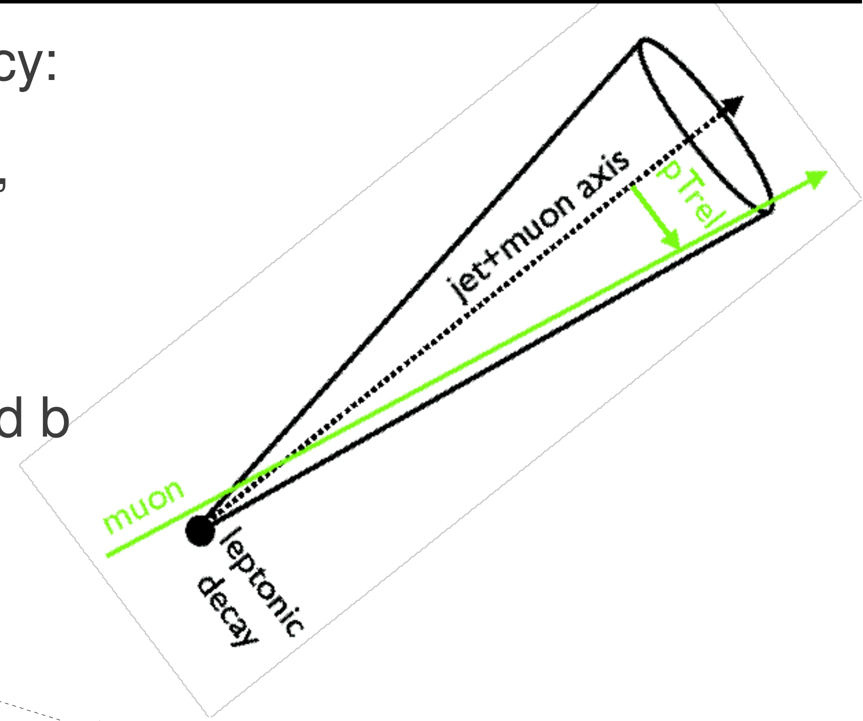


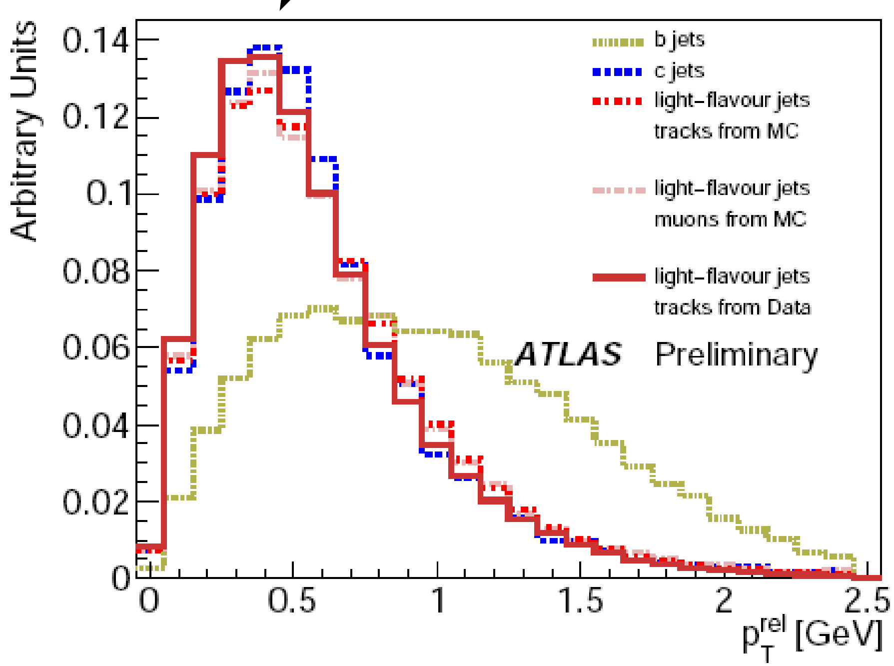
Combined algorithms: use information from both methods

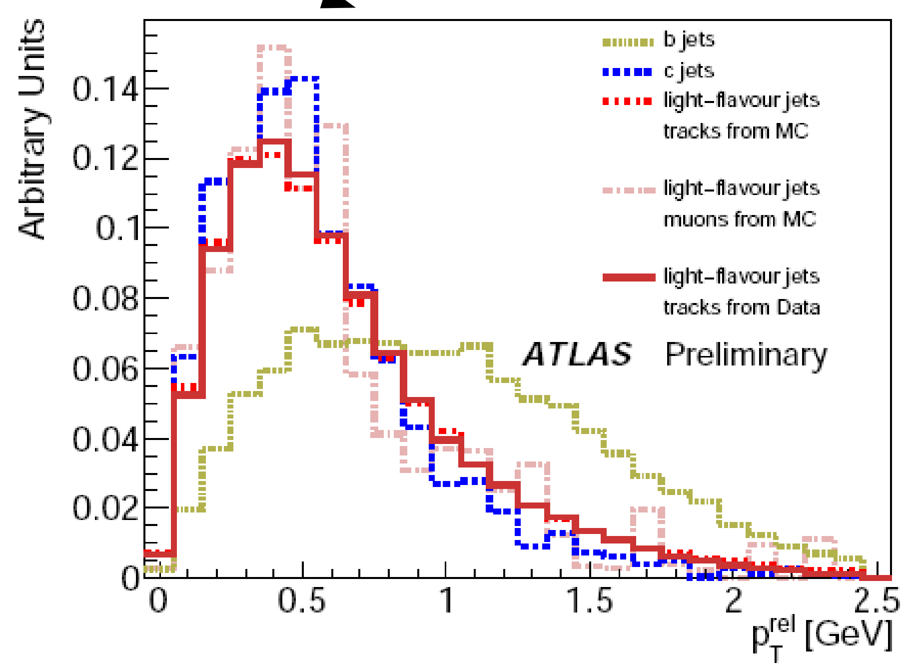
Calibration: the p_T^{rel} method

Data- driven method to calculate b-tagging efficiency:

- Select jets with muon in them → 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→ ratio of number of b-jets is the efficiency



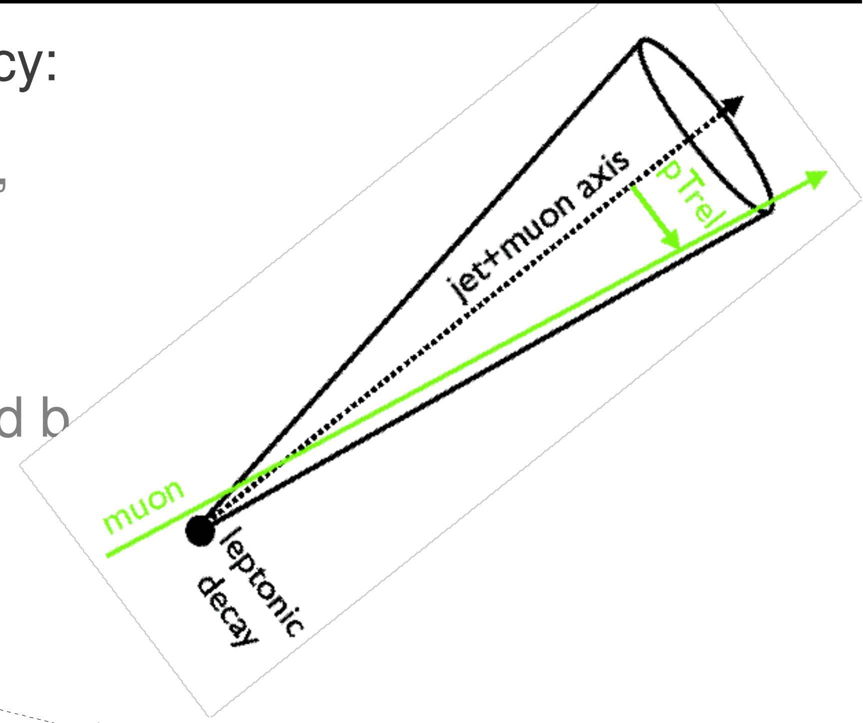


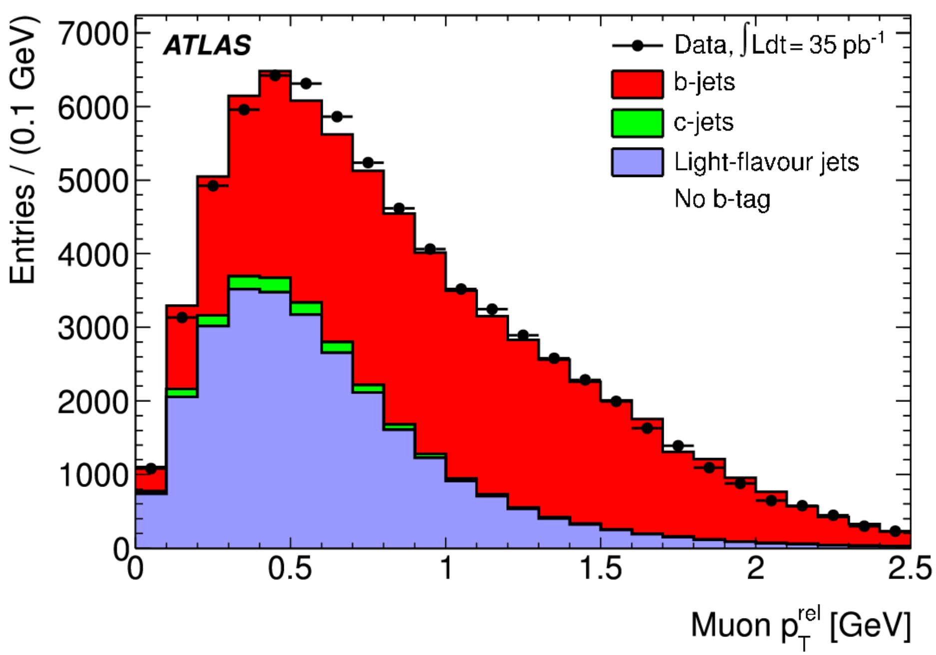


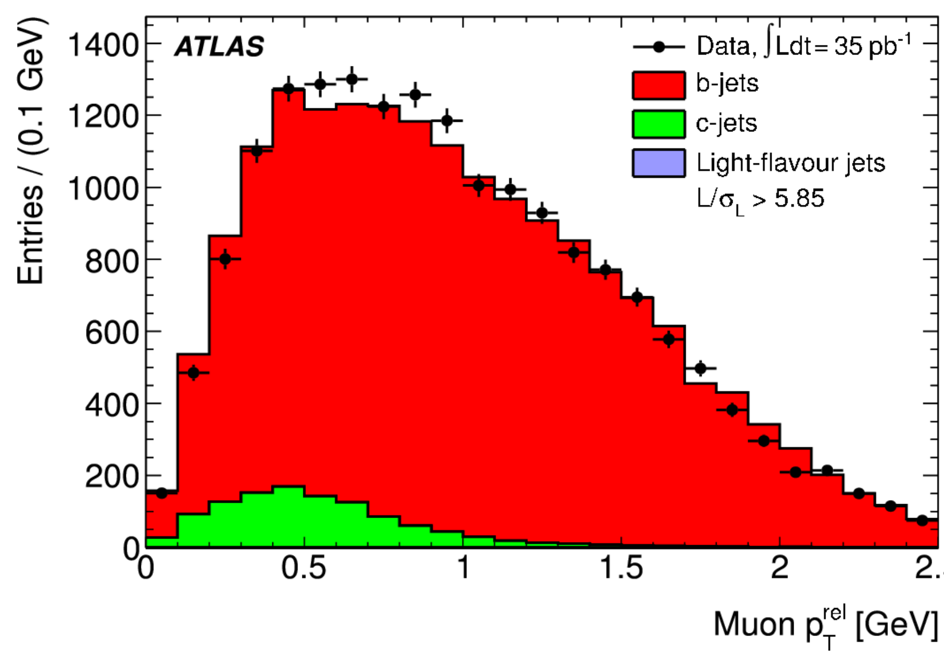
Calibration: the p_T^{rel} method

Data- driven method to calculate b-tagging efficiency:

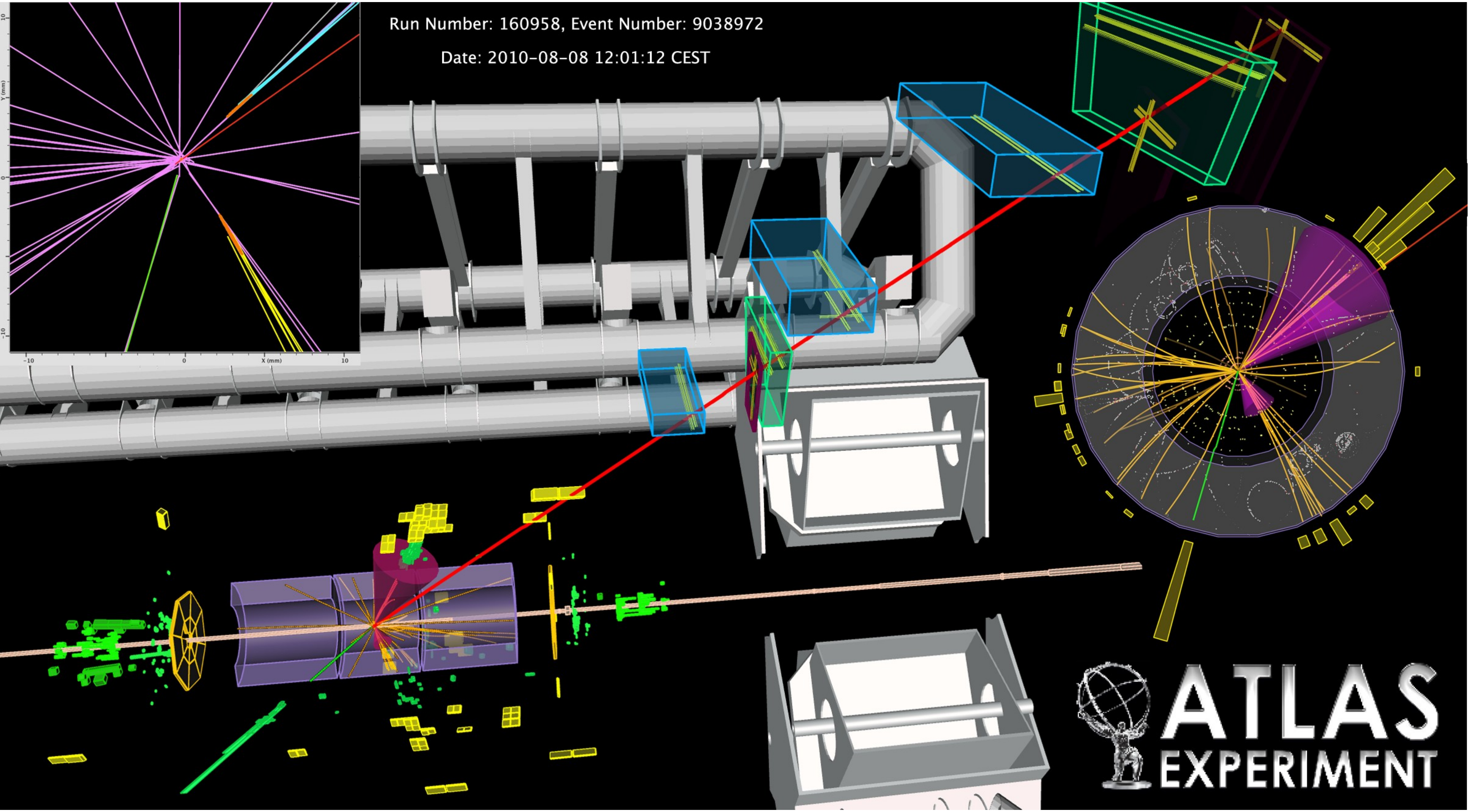
- Select jets with muon in them → enhanced in b, independent of track-based b-tagging
- Get template of relative p_T of the muon with respect to the jet axis (p_T^{rel}) for light (data), c and b (MC) jets, before and after SV0 cut
- Fit the data to them before and after SV0 cut→ ratio of number of b-jets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



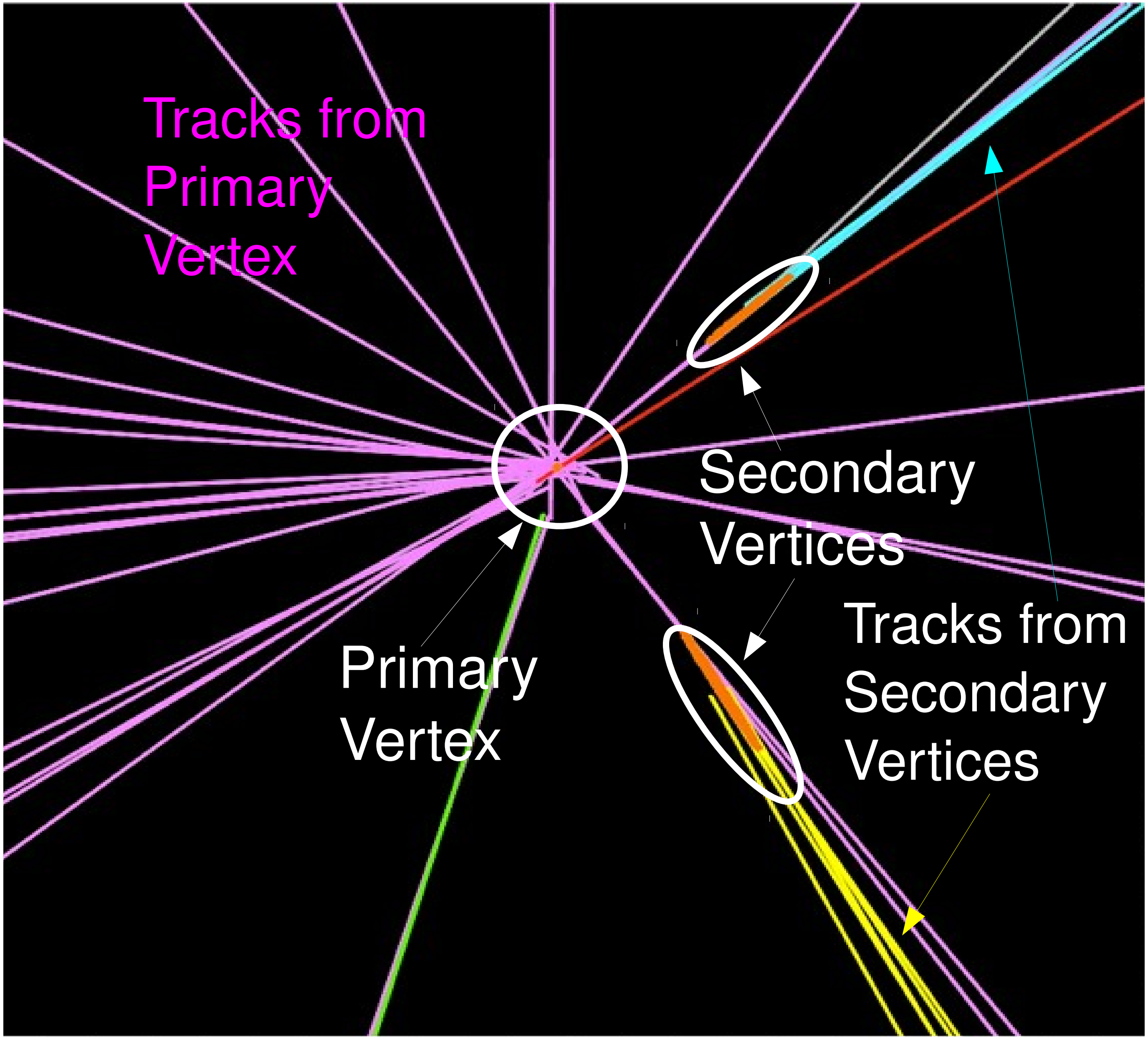




Track-based b-tagging



Track-based b-tagging



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

Inclusive b-jet and bb x-section

- 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 ~34 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

$$\frac{d^2\sigma_b}{dp_T dy} = \frac{1}{\Delta p_T \Delta y} \frac{N_b \cdot frac_b}{\varepsilon_{trig} \cdot \varepsilon_{sel} \cdot \varepsilon_{btag} \cdot L} \times C$$

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

Unfolding correction

Luminosity for the triggers used

Inclusive b-jet and bb x-section

- 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 ~34 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

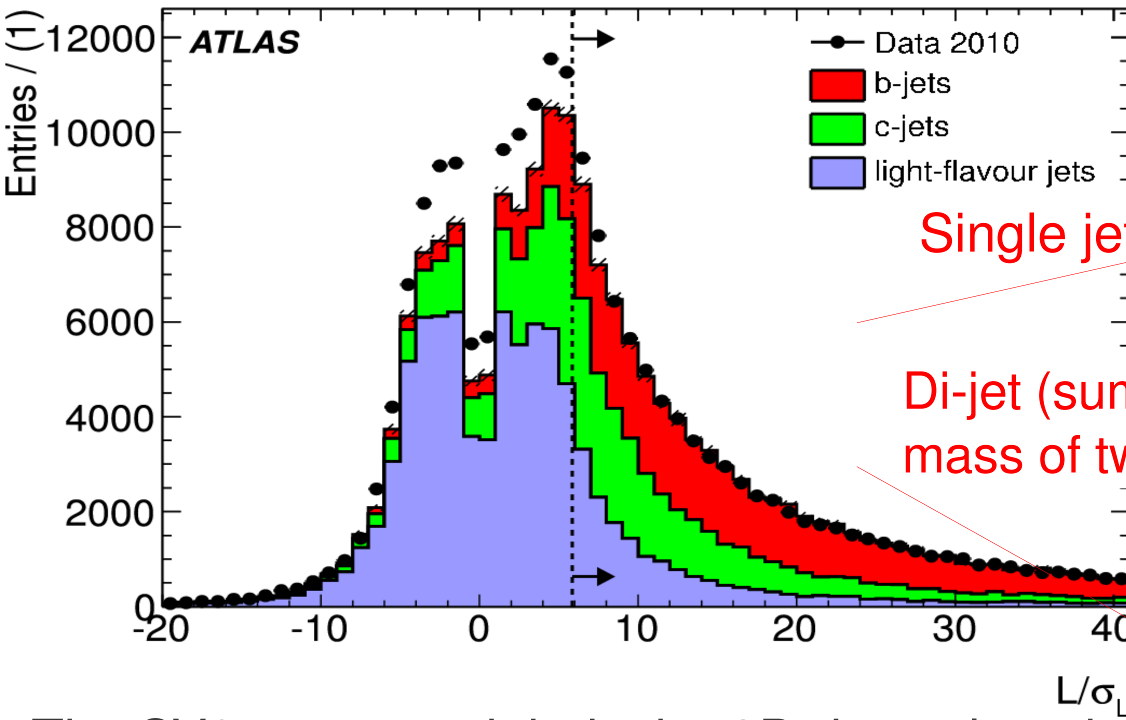
$$\frac{d\sigma_b}{dM} = \frac{1}{\Delta M} \frac{N_{b\bar{b}} \cdot frac_b}{\varepsilon_{trig}^{jj} \cdot \varepsilon_{sel}^{jj} \cdot \varepsilon_{btag}^{jj} \cdot L} \times C$$

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

Unfolding correction

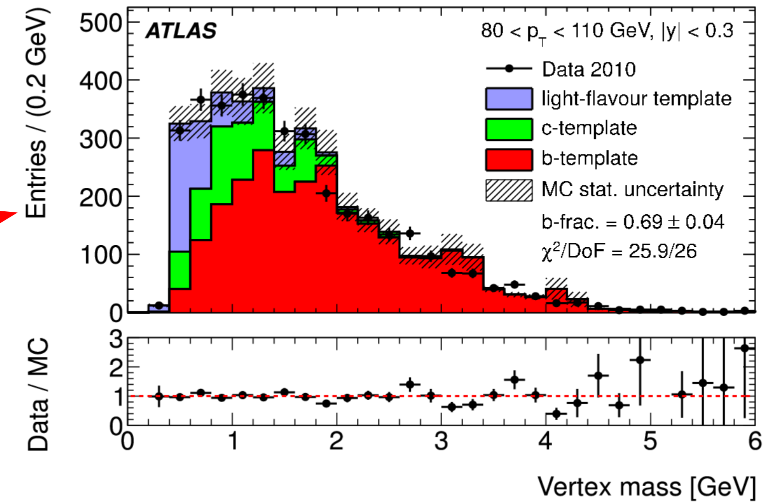
Luminosity for the triggers used

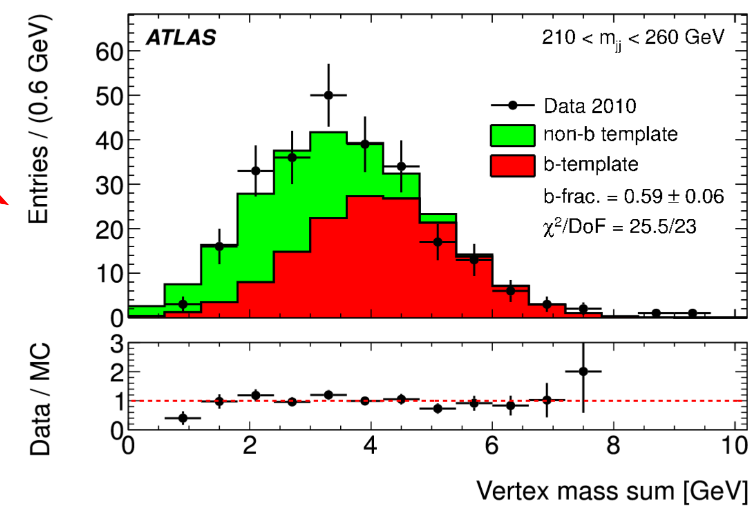
L/σ_L and SV mass template fit



Single jet

Di-jet (sum of SV mass of two jets)

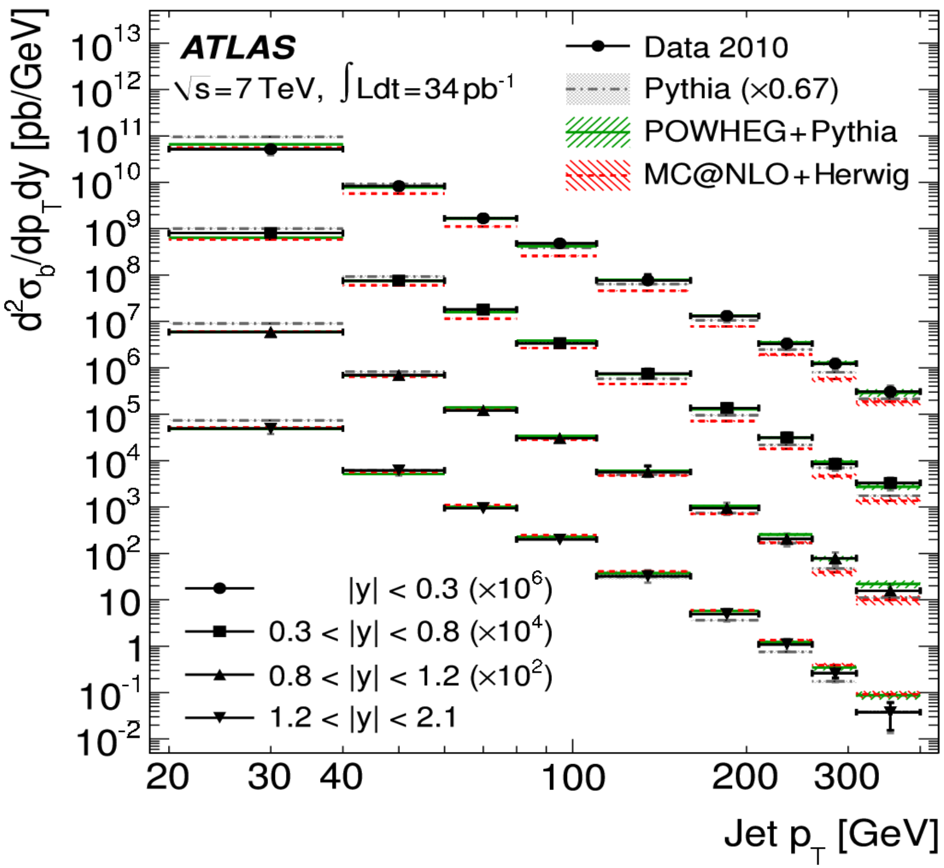


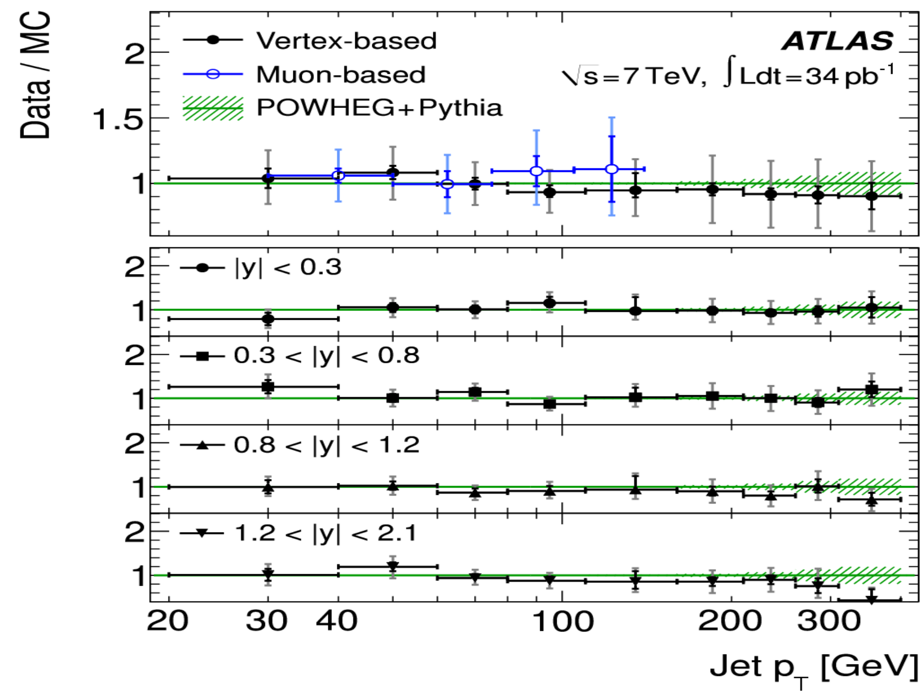


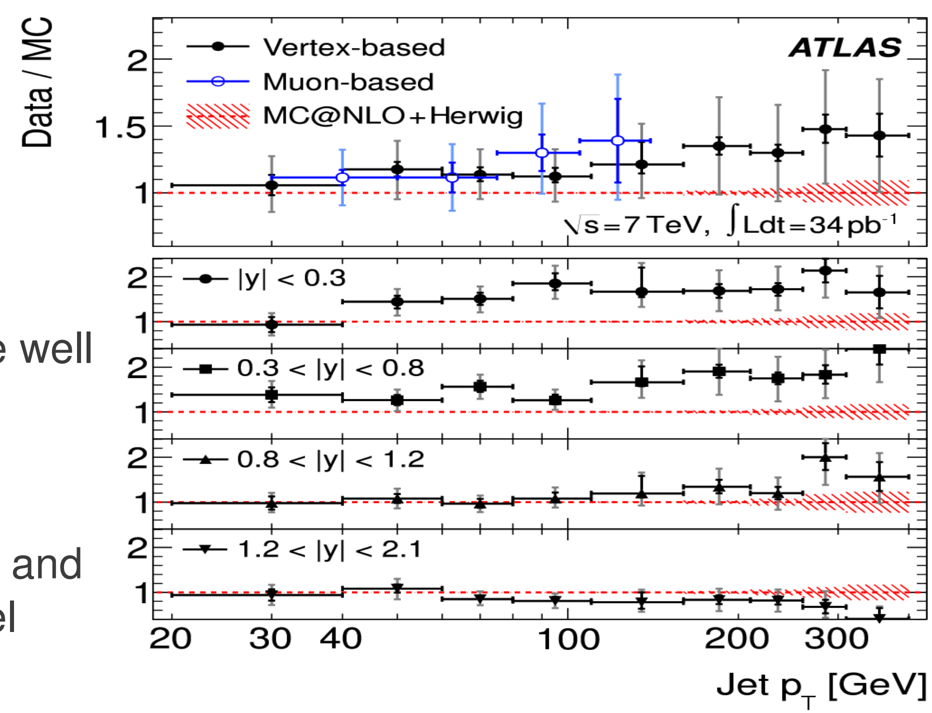
- The SV0 tagger weight is the 3D decay length significance signed w.r.t. the jet axis.
- Cut at 5.85 and use the invariant mass of the tracks in the secondary vertex to separate b- from non-b-jets in a template fit in each p_T-η or di-jet mass bin.
- Take light, charm and b (single-jet) or b and non-b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

Inclusive b cross-section



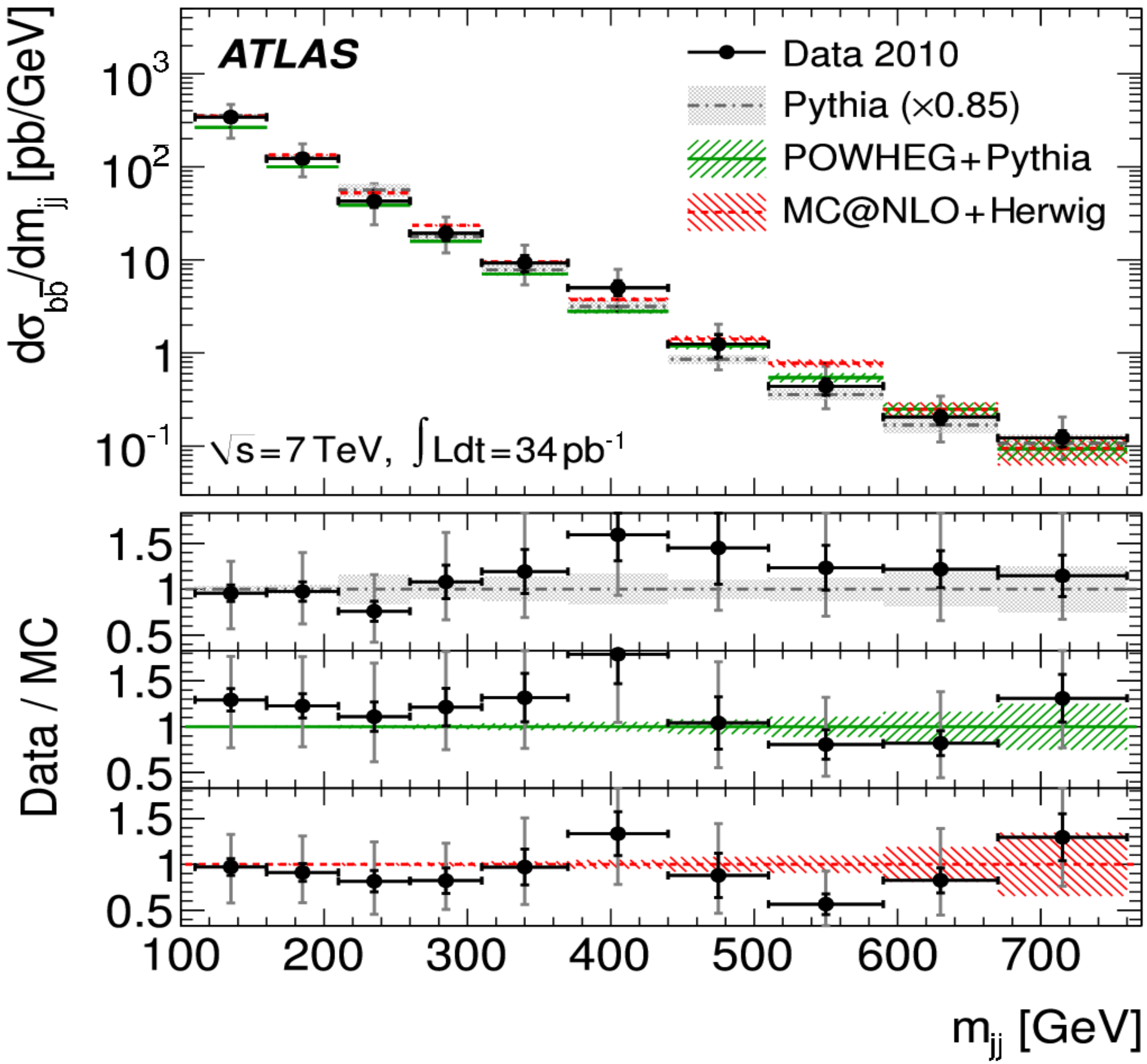




MC comparisons: broad agreement within the systematic uncertainties

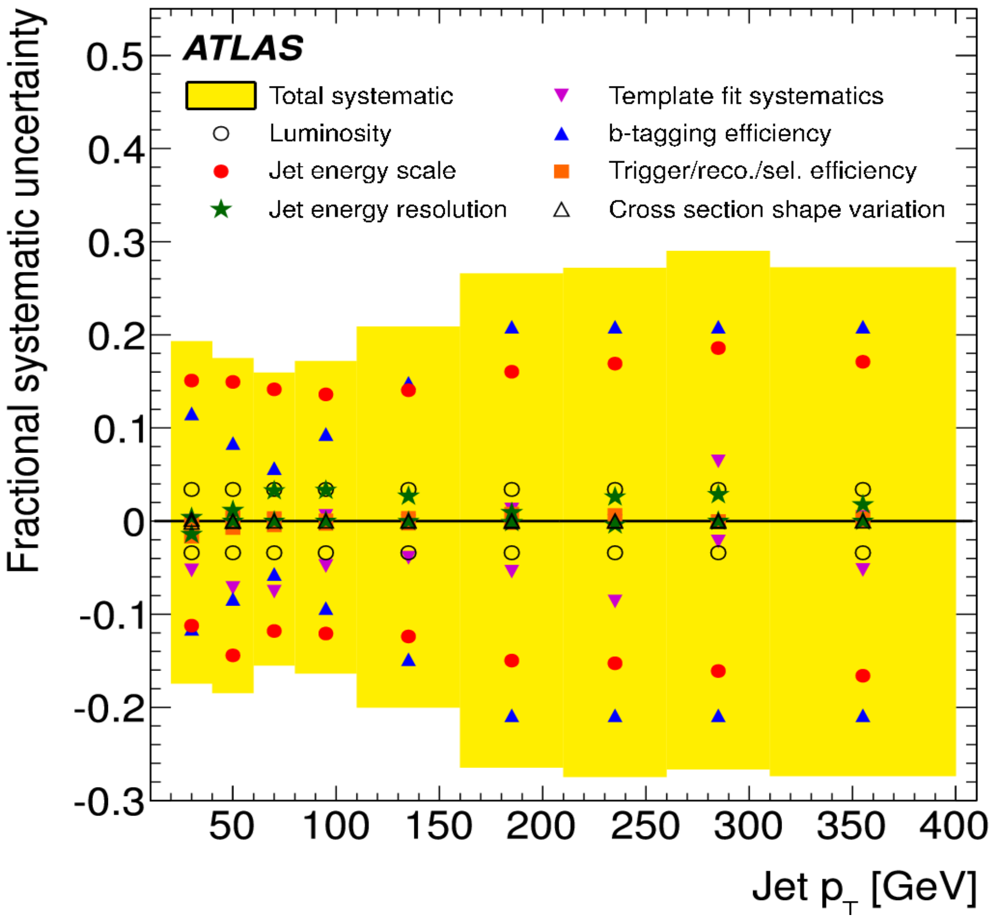
- Pythia6 (LL): normalization not expected to agree, shape well described
- POWHEG (NLO) + Pythia: good agreement
- MC@NLO + Herwig: disagreement for central rapidities and higher p_T , probably due to different hadronization model

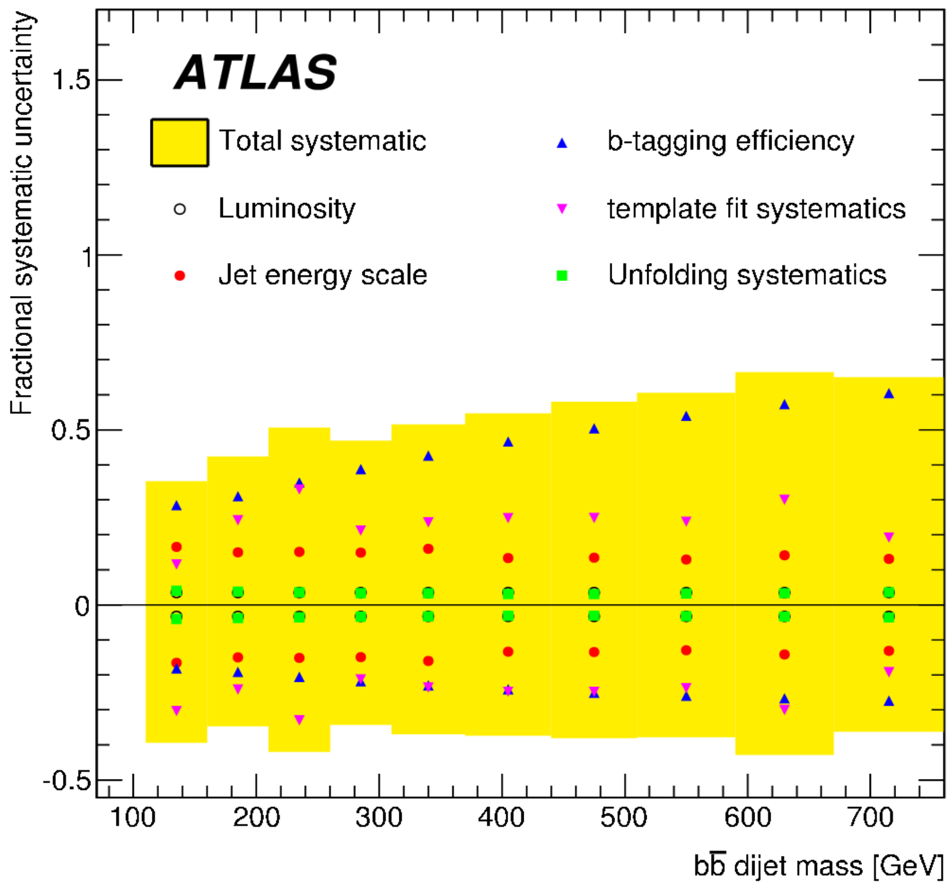
bb cross-section



Good agreement with Pythia, POWHEG and MC@NLO MC

Systematic uncertainties





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

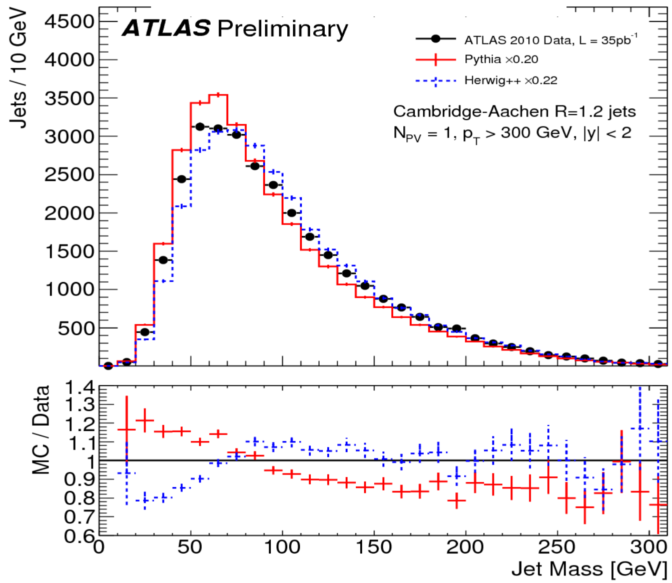
Jet mass and substructure

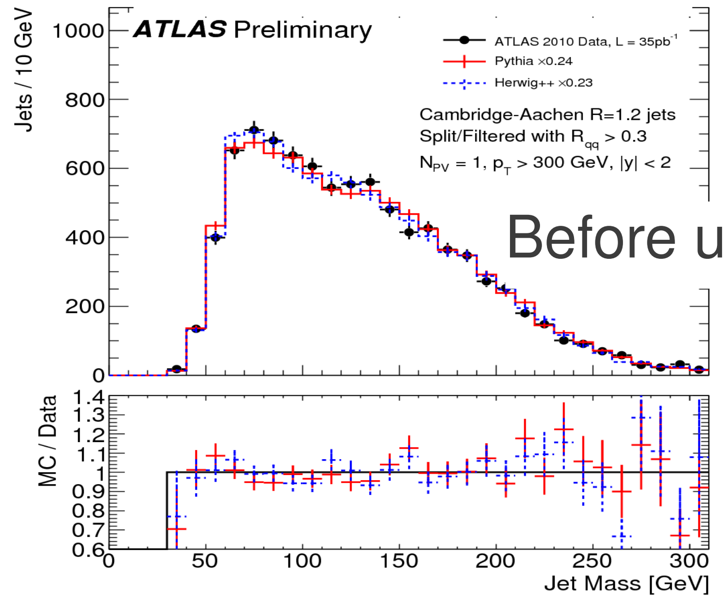
- Since we use jet mass and substructure, need to have confidence that MC describes such quantities well
- Show results for Cambridge-Aachen R=1.2 jets with p_T>300GeV and |y|<2, with 2010 data
- Dominant systematics: Jet Energy and Mass Scale, Jet Energy and Mass Resolution (non-standard algorithm,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

Jet mass distributions

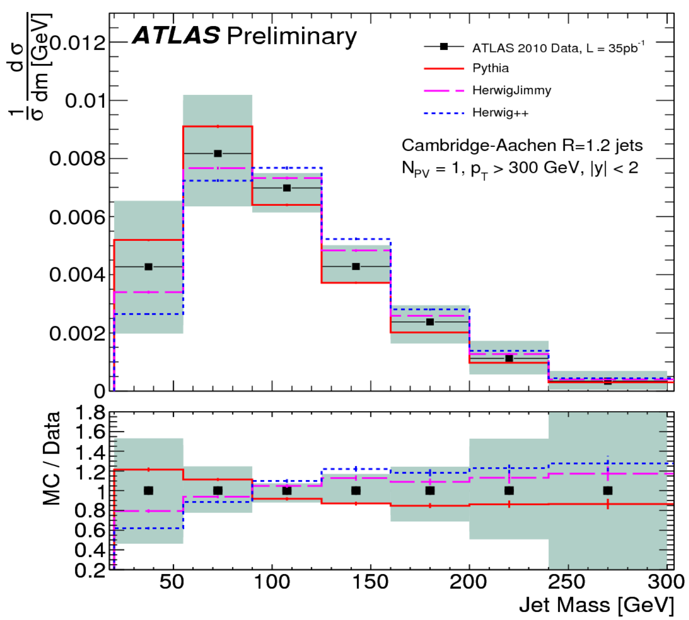
Original C-A jet

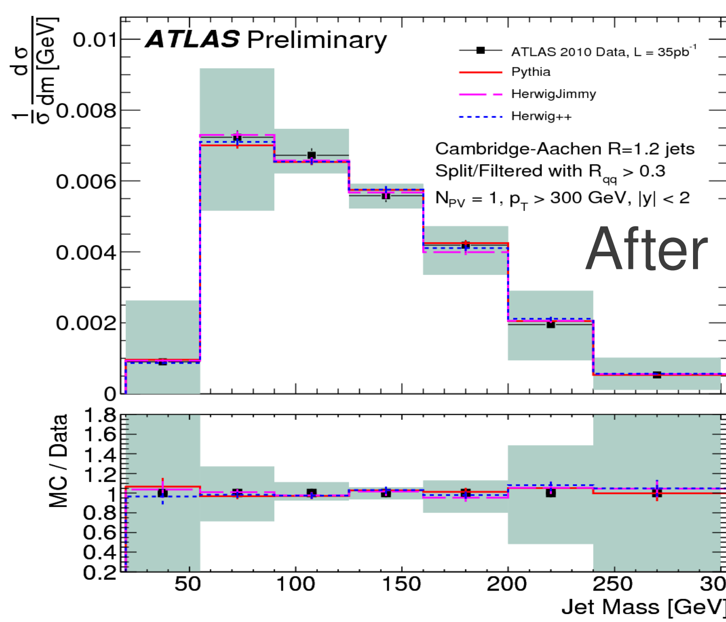
After splitting/filtering





Before unfolding (detector level)



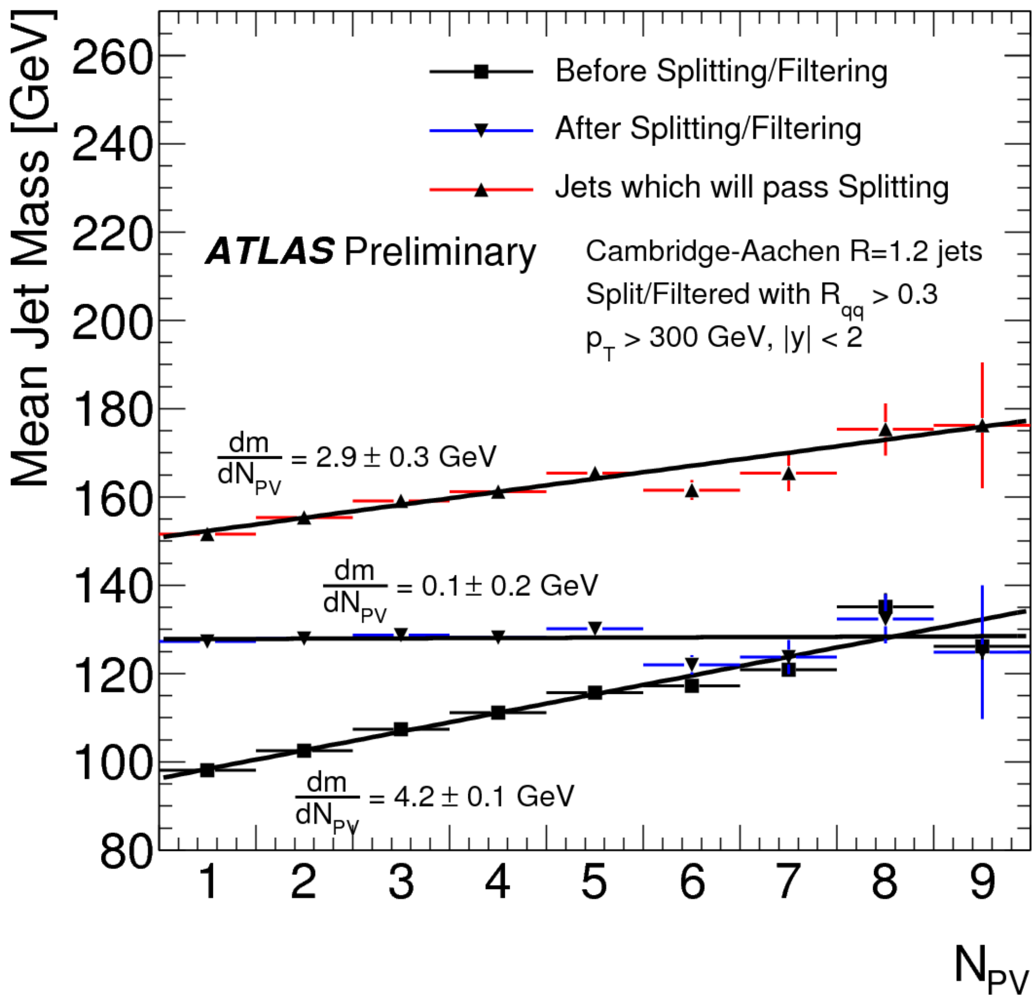


After unfolding (particle level)

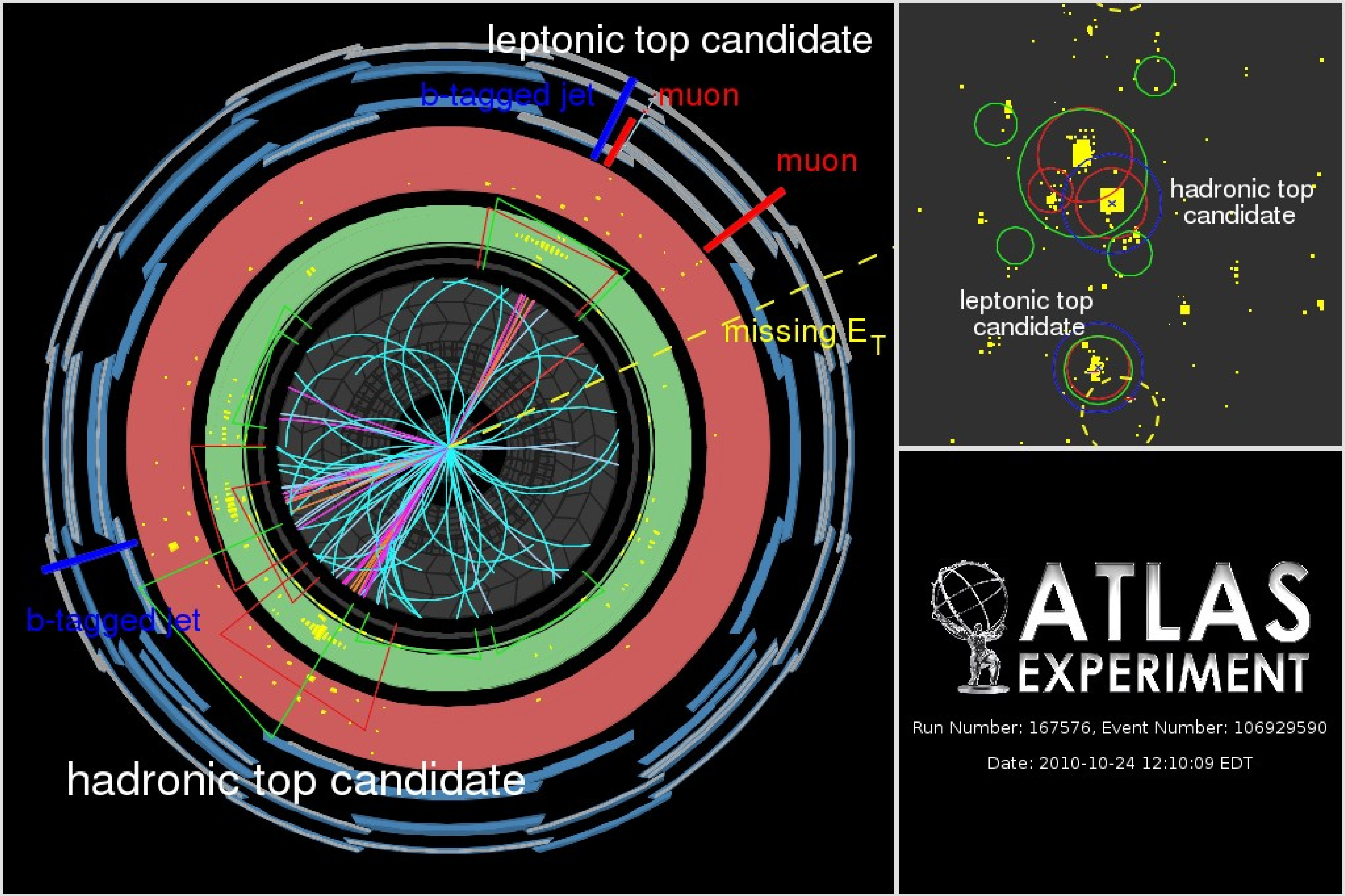
Good agreement with Pythia and Herwig, not so good with Herwig++ (too many high mass jets)

Pileup

- 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.



Jet with substructure

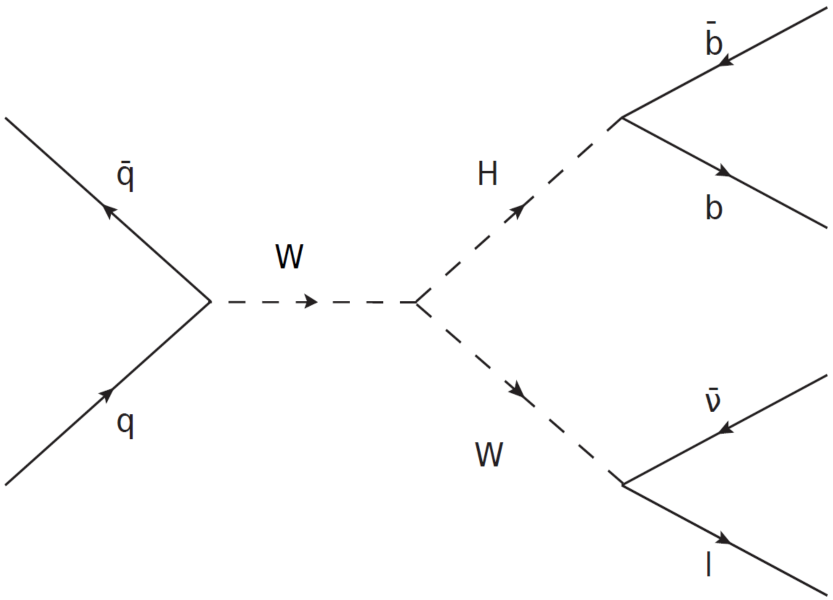


Red – antiKt04, blue – antiKt06, green – antiKt1

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

Inclusive VH→bb

- 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 ($W \to l\nu$, $Z \to ll$, $l = e,\mu$)
- Look for two additional b-tagged jets in the event and reconstruct the Higgs mass as m_{bb}
- Backgrounds: Vbb (irreducible), ttbar, diboson (irreducible), QCD



Selection

WH → bb (in general tighter)

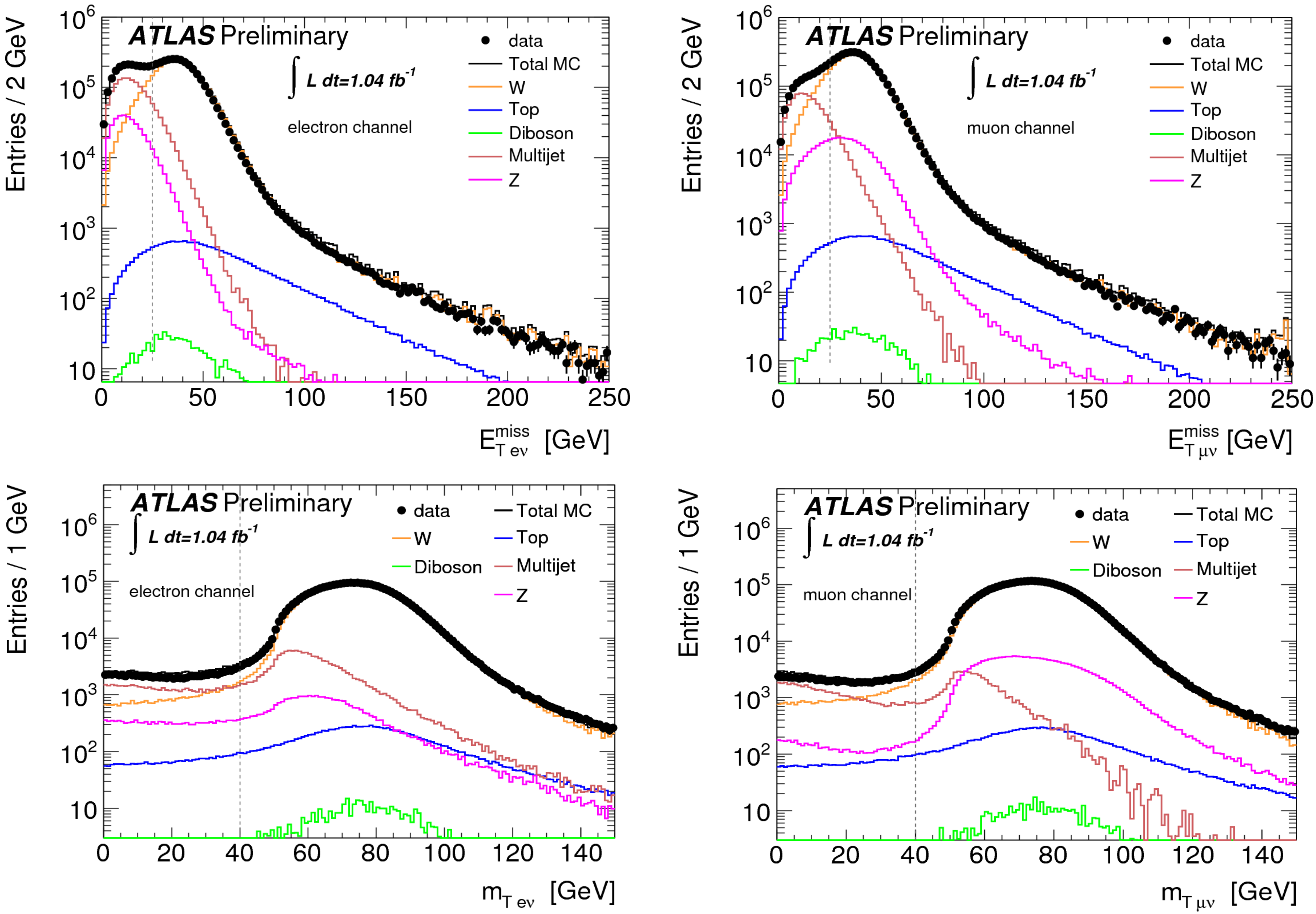
- One good lepton in the event
- E_T^{miss} > 25GeV, m_T>40GeV
- Exactly two jets in the event, both b-tagged – to reduce ttbar

ZH → bb

- Two good leptons in the event at m_Z
- E_T^{miss} < 50GeV
- At least two jets in the event, the two leading ones b-tagged

- 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

Some distributions...

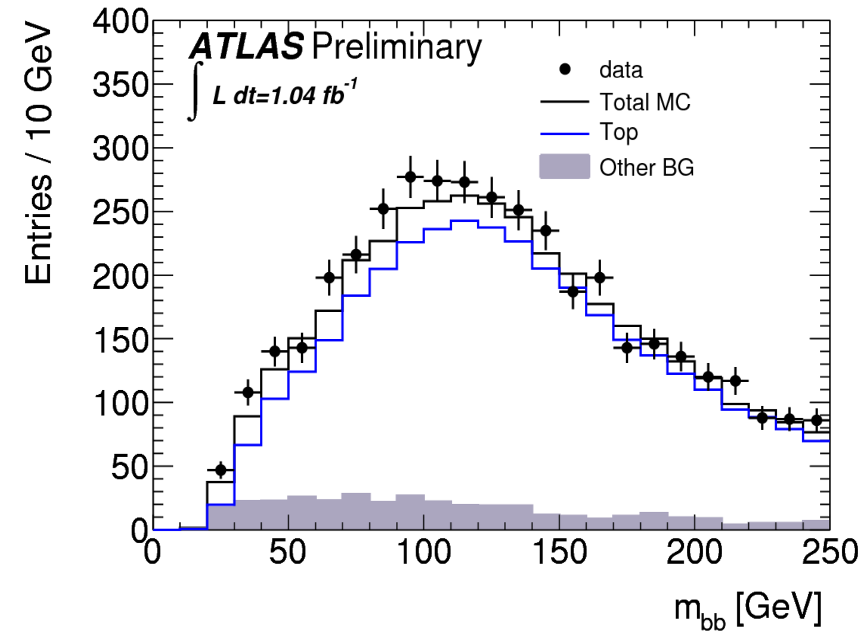


WH background estimations

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

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

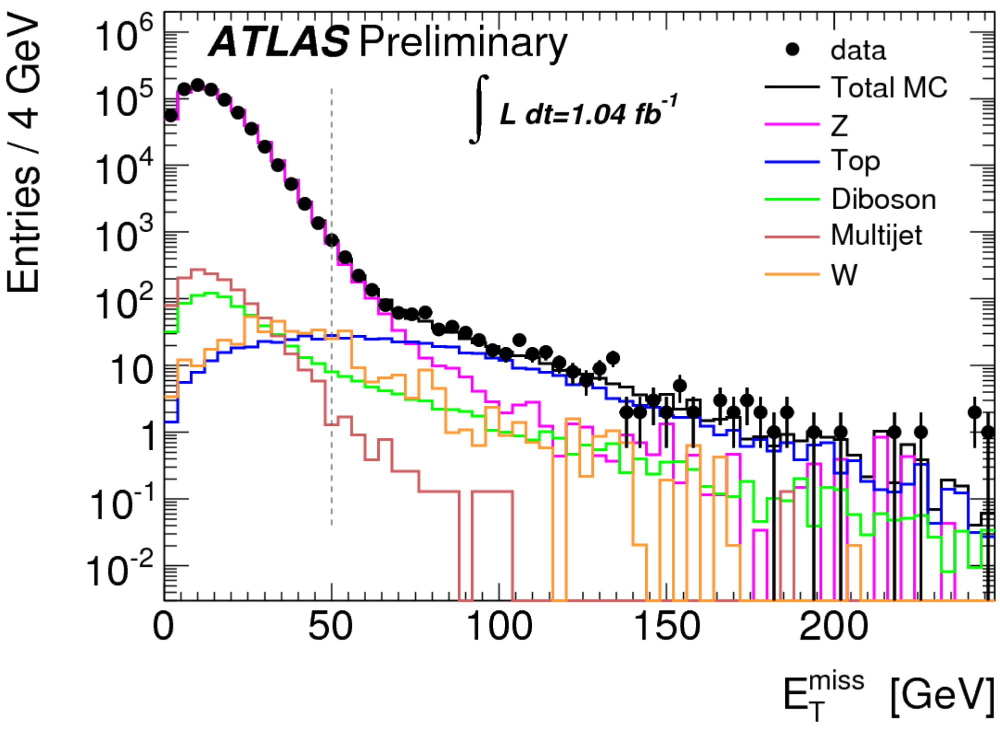


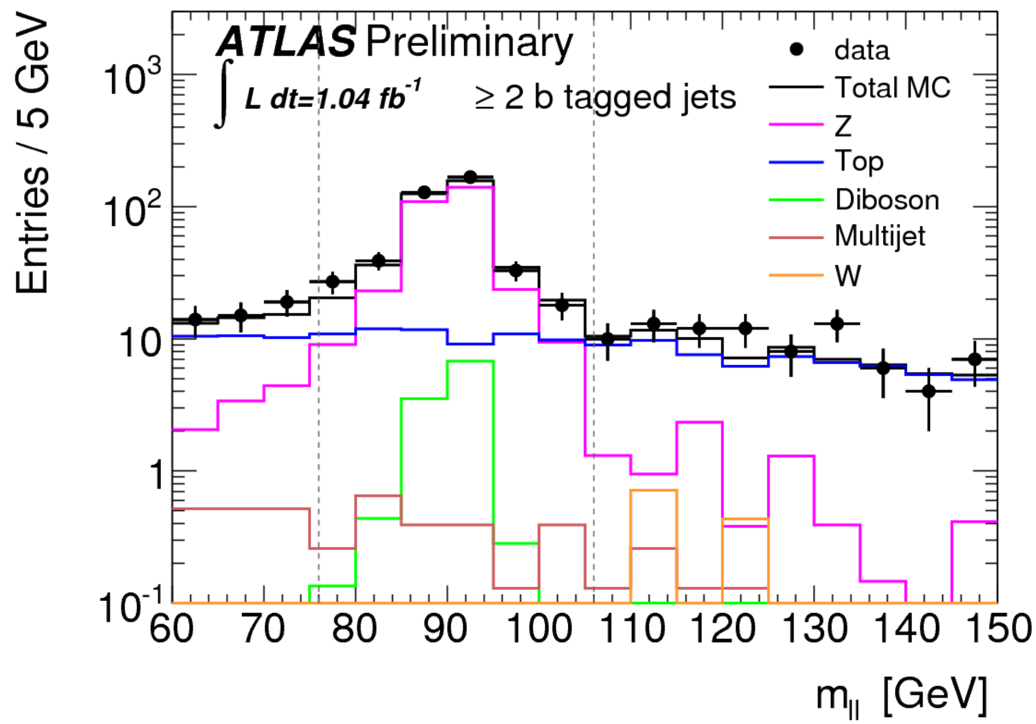
WZ and WW: irreducible background. Too small to measure from data, use MC

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.

ZH

Some distributions...

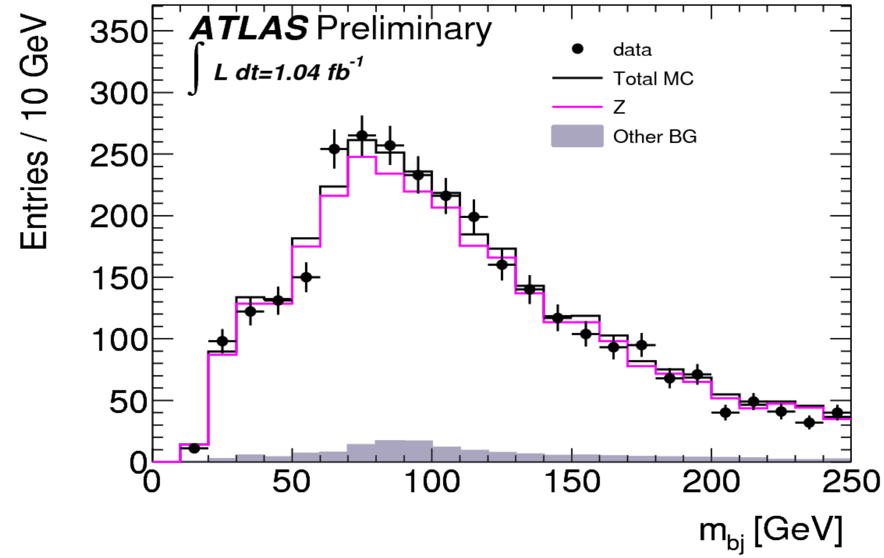


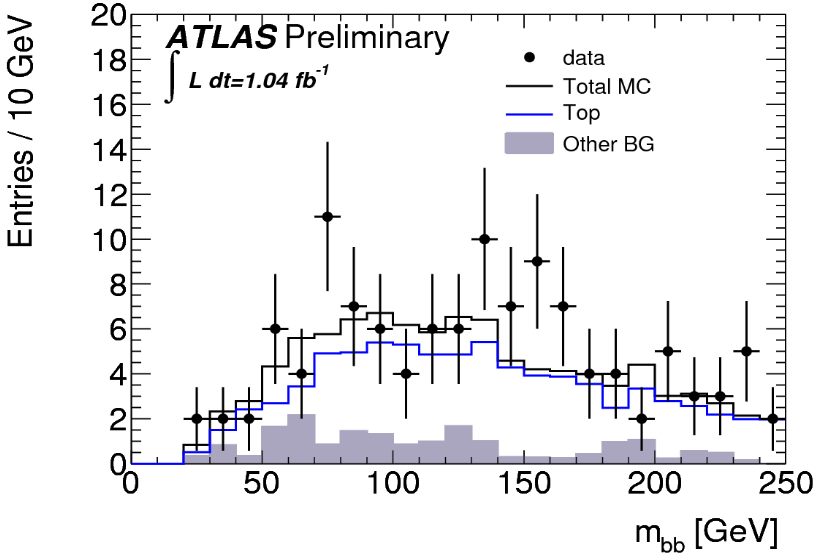


ZH background estimations

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

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





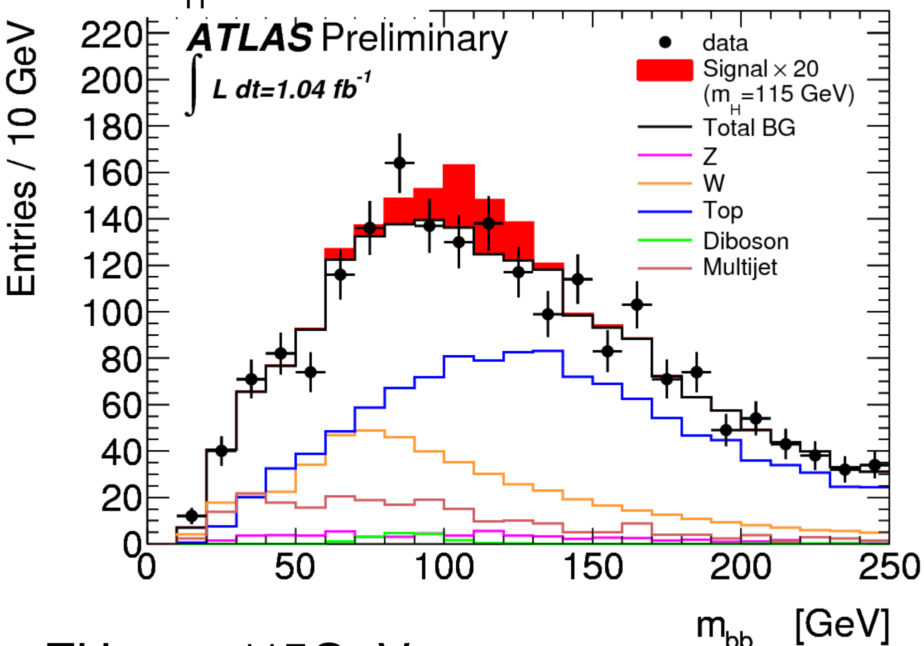
Top (ttbar and single): use MC, verify normalization in data using m_{ll} sidebands (60GeV<m_{ll}<76GeV and 106GeV<m_{ll}<150GeV)

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

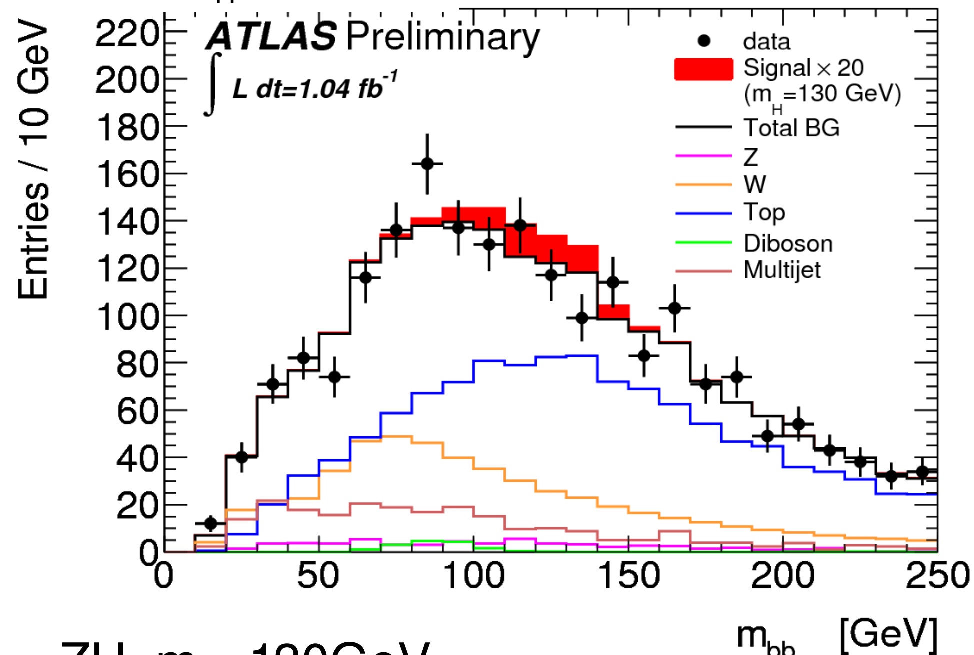
QCD multijet (fake or true leptons making a "Z"): template method for e, MC for μ (negligible). Derive templates from data enriched in fake electrons, fit di-lepton mass distribution and use normalization.

Results - m_{bb}

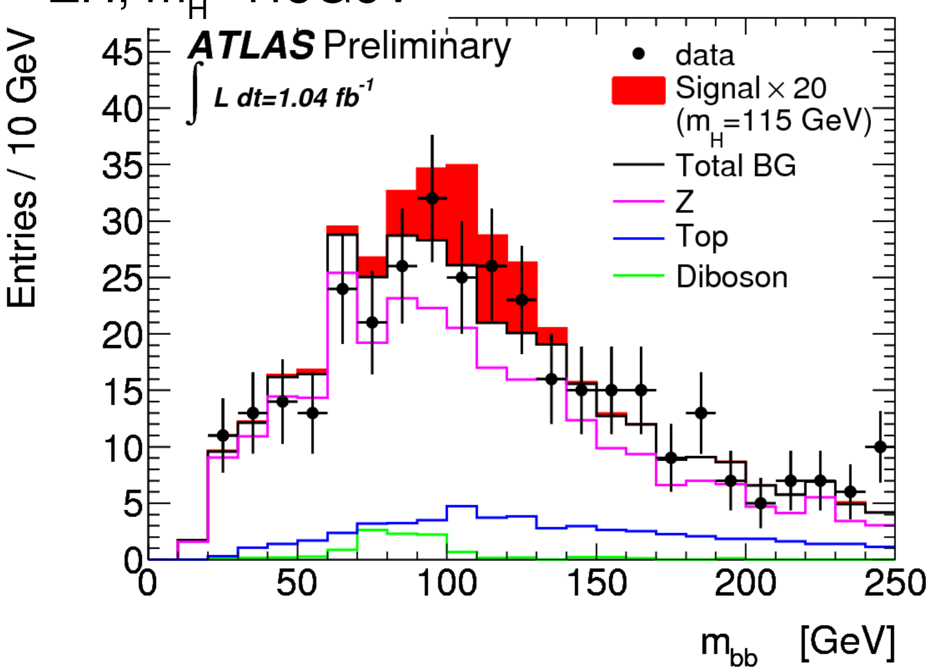
WH, m_H=115GeV



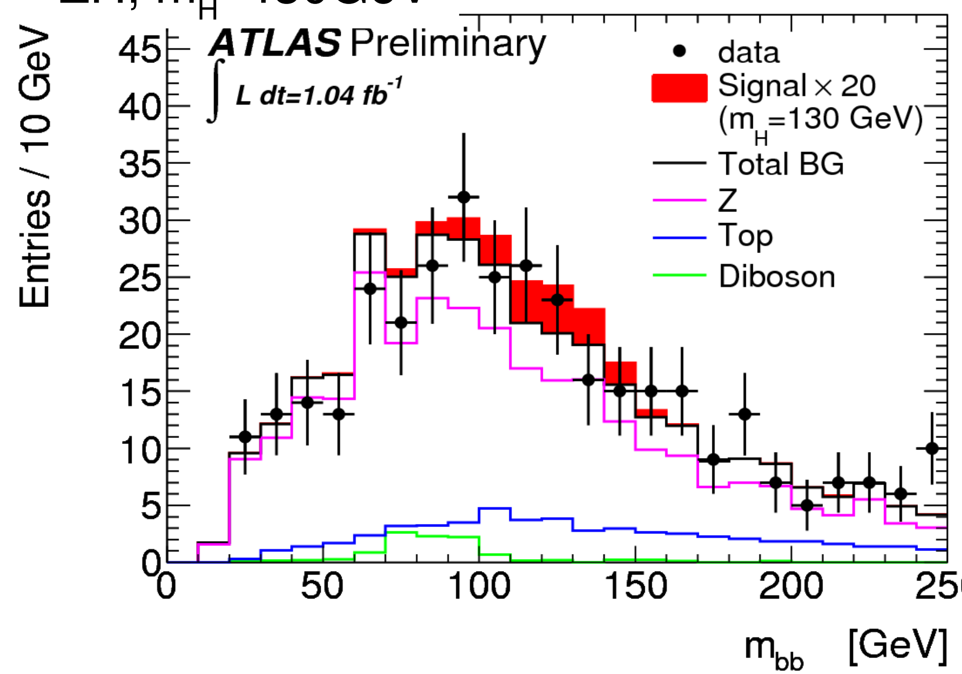
WH, m_H=130GeV



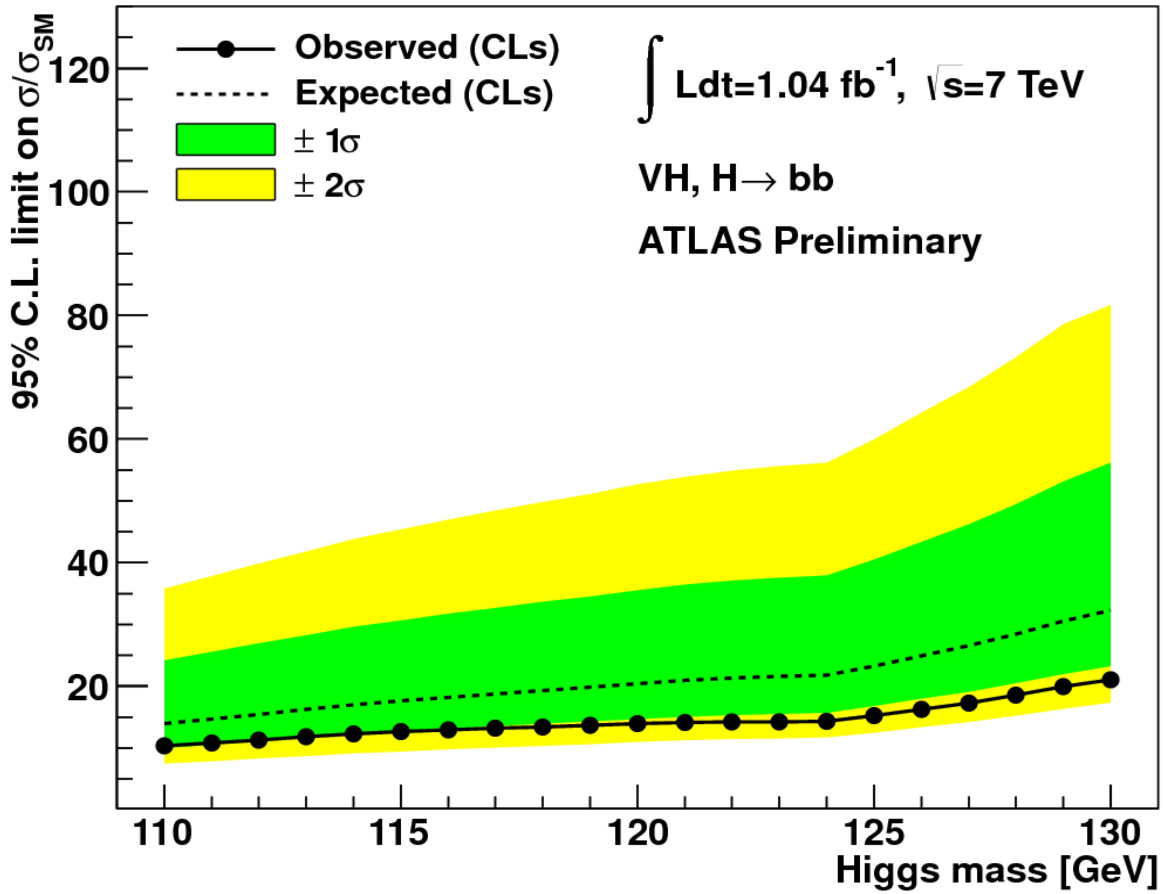
ZH, m_H=115GeV



ZH, m_H=130GeV



Results - limits



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 (σ_{mH}~18-20GeV), dominated by out-of-cone losses

Systematics

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

Source of Uncertainty	Treatment in analysis	
	ZH	*WH*
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 $ZH \to \ell\ell b\bar{b}$ signal		Effect on $WH \to \ell\nu b\bar{b}$ signal	
	m_H = 115 GeV	m_H = 130 GeV	m_H = 115 GeV	m_H = 130 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%
b-tagging Efficiency	16%	17%	16%	17%
b-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→bb ?
- Track-based b-tagging with ATLAS
- Associated analyses of a boosted VH→bb search
  - Inclusive b-jet and bb cross section
  - Jet substructure measurements
  - Un-boosted (inclusive) VH→bb
- Conclusions

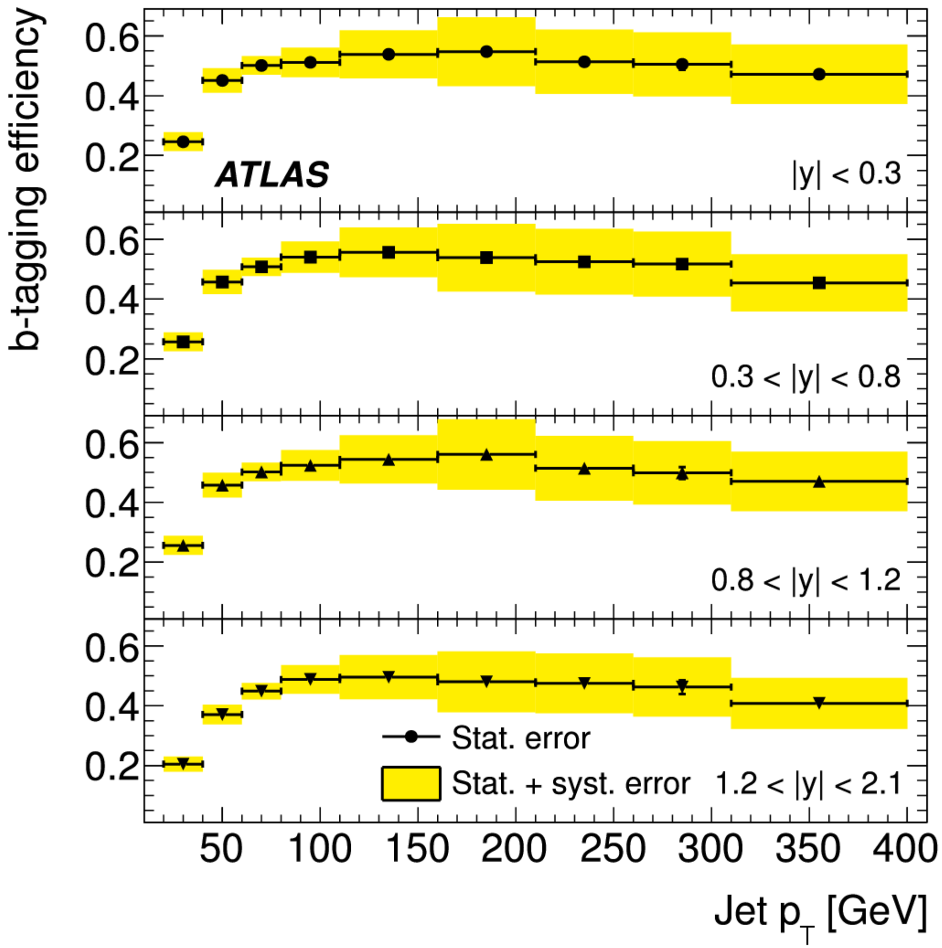
Towards a boosted VH→bb

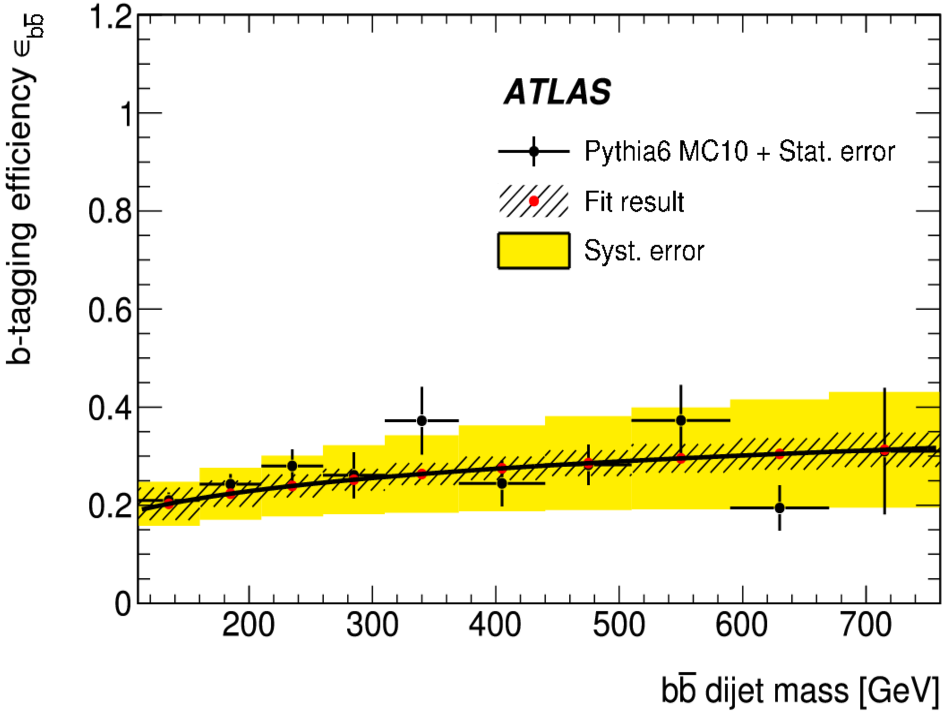
- Boosted VH→bb can contribute valuable sensitivity to low-mass Higgs searches
- ATLAS has proven good understanding of jets, b-tagging and even jet substructure, as shown by early SM and performance measurements
- Inclusive VH→bb measurement available, working on further improvements
- Moving on to boosted VH→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→bb to constrain some of those systematics

EXTRAS

B-tagging efficiency

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





Selection

WH → bb

- Electrons: "tight" quality, p_T>25GeV, |eta|<2.47, isolation, $|d_o|$<0.1mm, $|z_o|$<10mm
- Muons: p_T>25GeV, |eta|<2.4, isolation, $|d_o|$<0.1mm, $|z_o|$<10mm, ΔR>0.4 from jets
- One lepton in the event
- E_T^{miss} > 25GeV, m_T>40GeV
- Exactly two jets in the event, both b-tagged

ZH → bb

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

- Jets (antiKt algorithm with ΔR=0.4): jet cleaning, p_T>25GeV, Jet Vertex Fraction > 0.75, |eta|<2.5, ΔR>0.4 from electrons
- B-tagging: SV+IP3D combined tagger, at 70% efficiency – 50 light rejection working point on ttbar simulation
- Single-lepton and di-electron (Z channel) triggers
- Primary vertex with at least 3 tracks