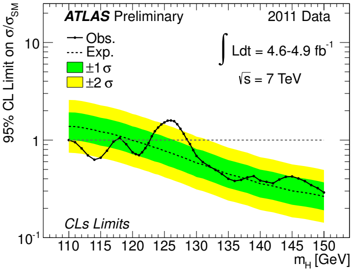
Boosted $H \rightarrow b\bar{b}$ analysis

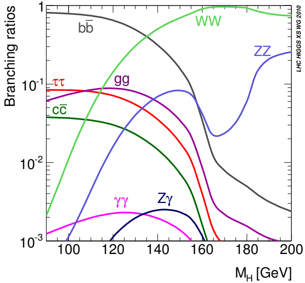
Brendan O'Brien - University of Edinburgh

April 2 - 4, 2012

Why $H \to b\bar{b}$?

- Present fits to data, including exclusions from LEP, Tevatron, and LHC predict a Higgs boson in a small mass window.
- In this range, considered to be "low mass", the dominant decay branching ratio is for the $H \to b\bar{b}$ process.
- The most promising production method is via $q\bar{q} \to VH$.





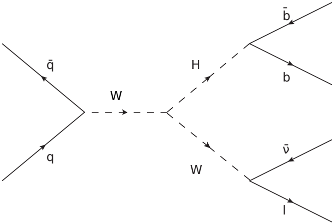
[1]

Why boosted $H \to b\bar{b}$?

- The $H \to b\bar{b}$ process will produce two b-jets, which must be distinguished from other background processes.
- Many background processes exist that also produce two b-jets.
- Two b-jets resulting from a high P_T Higgs ($\gtrsim$ 200 GeV) will be close together, forming a so-called fat-jet.
- This b-jet closeness requirement allows powerful separation from background processes, at the expense of $\sim$ 95% of the signal.

Analysis overview

- Start with $q\bar{q} \to VH$ production.
- Find leptonically decaying vector-boson (e or μ).
- Find hadronically decaying Higgs ($H \to b\bar{b}$).

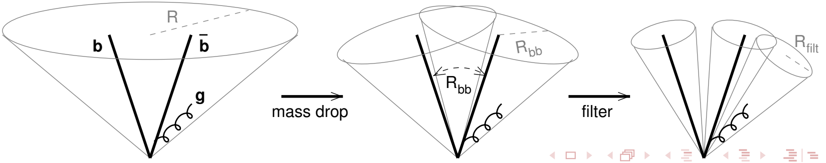


$q\bar{q} \to WH \to b\bar{b}l\bar{\nu}$ process

- Select events with large Higgs P_T, and a Cambridge-Aachen (C-A) reconstructed fat-jet containing both b-jets.
- Use substructure techniques to split the fat-jet into subjets and determine their invariant mass.

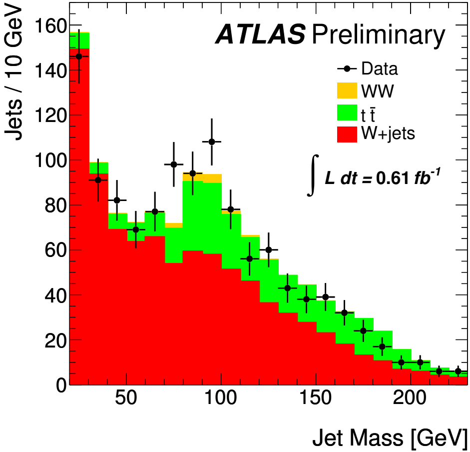
Working with jet substructure

- Jets are reconstructed using C-A algorithm (R = 1.2).
- Often ATLAS jets are clustered using inverse P_T ($AntiK_T$). C-A jets are clustered using physical distance only.
- Clustering history is searched in reverse until a point where a significant mass drop is identified ($> \frac{1}{3}$).
- Smaller R used to re-cluster the remaining constituents, filtering out the irrelevant radiation.
- Three highest P_T subjets used to form Higgs candidate.



Jet mass distribution

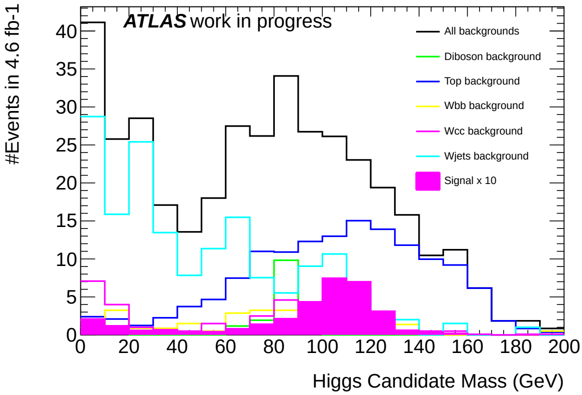
- Result shown in conference in 2011 [2].
- Mass distribution for split and filtered jets with no *b*-tagging applied.



Selection for $WH \rightarrow l\nu b\bar{b}$

- Standard e, μ, and jet selections following ATLAS recommendations (see backup slides), then:
- One charged lepton with $P_T > 20$ GeV and $E_T^{miss} > 25$ GeV
- W candidate with $M_T > 40$ GeV
- One fat-jet with $P_T > 180$ GeV, containing at least two subjets.
- Veto events with any $AntiK_T$ jets with $P_T > 20$ GeV and not within 1.2 of the signal C-A jet.
- $\Delta\phi_{W,H} > \frac{2\pi}{3}$
- W candidate with $P_T > 200$ GeV
- b-tagging: MV1 (Multi Variate tagger) weight > 0.60 (70% efficiency).

M_H with ≥ 1 b-tag plot



- $M_H = 120$ GeV

M_H with 2 b-tag plot

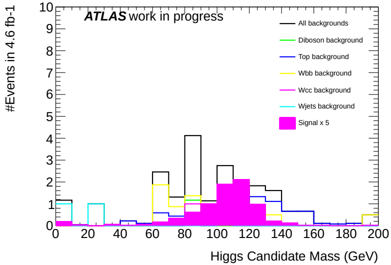


Figure: Boosted analysis with 2 b-tags

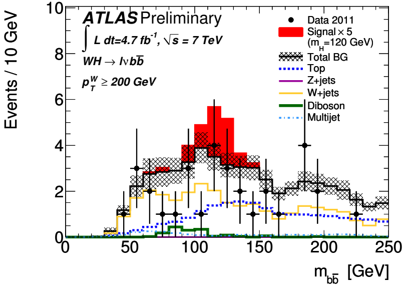


Figure: Unboosted analysis in highest P_T bin [3]

Current status

- All results shown here are a work in progress, and reflect continuing efforts in this analysis.
- Shown for boosted $WH \rightarrow l\nu b\bar{b}$ only. $ZH \rightarrow l^+l^-b\bar{b}$ and $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ analyses are being developed also, but less mature.
- Boosted $Z \rightarrow b\bar{b}$ analysis is also ongoing in ATLAS, expecting observation and constraints on WH/ZH systematics.
- Working on neural network, and boosted decision tree approach in parallel. Expected to improve the $\frac{S}{\sqrt{B}}$.
- Planning data-driven background determination.

ΔR dependence of b-tagging

- Data/MC scale factors are required for b-tagging efficiency in subjets.
- Current b-tagging calibrations are derived based on $AntiK_T$ jets with R $= 0.4$. These do agree well with C-A jets in the $\Delta R > 0.4$ regime.
- But many subjets have $\Delta R < 0.4$. Closer jets have more ambiguous track-jet matching and are less efficient to b-tag .
- Need to validate b-tagging in low ΔR regime with subjets, in order to use it in this (and other) analysis involving boosted jets.

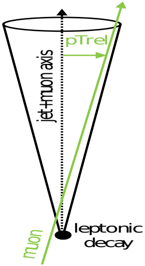
ΔR dependence of b-tagging

- Data/MC comparisons have been done on b-tagging relevant quantities in low and high ΔR regimes. Agreement reasonable.
- Some data/MC discrepancies, but the integral b-tagging weight has agreement within 5%.
- Quantify extra systematic for $\Delta R < 0.4$ region in the short term, derive ΔR dependent scale factors with the P_T^{Rel} method in the long term.

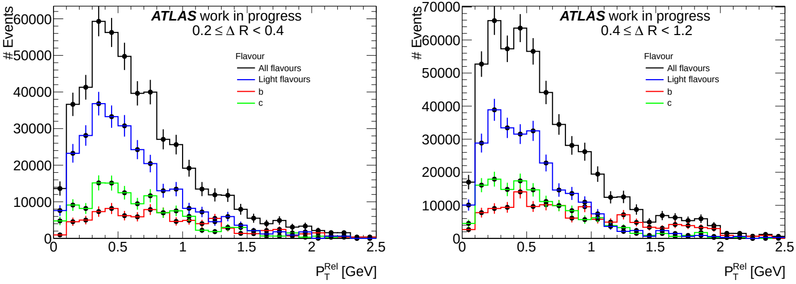
P_T^{Rel} method

- P_T^{Rel} method can be used to obtain data/MC scale factors for the b-tagging efficiency. So far this has only been made use of for $AntiK_T$ jets R = 0.4.

- P_T^{Rel} makes use of flavour templates obtained by investigating the momentum of muons, transverse to muon-jet axis in leptonic b decays.
- Investigations into the applicability of using the P_T^{Rel} method for subjets are underway.



P_T^{Rel} method



- Above studies used pythia jetjet monte carlo.
- Preliminary templates have been derived for various ΔR ranges, and discriminating power between flavours looks reasonable even at low ΔR.
- More investigations and detailed template derivation needed.

Outlook

Boosted substructure analysis is a promising technique, and the analysis is becoming more mature.

- Boosted $H \rightarrow b\bar{b}$ analysis:
  - Optimize cuts. and work towards $\frac{S}{\sqrt{B}}$ improvement.
  - Evaluate substructure systematics.
  - Aim to combine boosted analysis with unboosted analysis in highest P_T bin for summer conferences.
- Boosted b-tagging related:
  - Use current P_T^{Rel} derived scale factors with the subjets, but inflate systematic for low ΔR region (< 0.4), for short term.
  - Explicitly derive ΔR dependent P_T^{Rel} calibration for C-A subjets, $\sim$ 6 months.

References

An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb^{-1} of *pp* collision data at $\sqrt{s} = 7$ TeV.
Technical Report ATLAS-CONF-2012-019, CERN, Geneva, Mar 2012.

Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a *b*-quark pair with the ATLAS detector at the LHC.
Technical Report ATLAS-CONF-2011-103, CERN, Geneva, Jul 2011.

Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a *b*-quark pair using up to 4.7 fb^{-1} of *pp* collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC.
Technical Report ATLAS-CONF-2012-015, CERN, Geneva, Mar 2012.

Backup slides

Lepton selection

- Electrons
  - $E_T > 20$ GeV
  - $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
  - Author 1 || 3, isEM && track match, $\frac{\sum E_T}{E_T} < 0.14$, $\frac{\sum P_T}{P_T} < 0.1$
- Muons
  - Staco || Muid algorithms
  - $P_T > 20$ GeV
  - $|\eta| < 2.7$
  - $\frac{\sum E_T}{E_T} < 0.14$, $\frac{\sum P_T}{P_T} < 0.1$

Jet selection

- *AntiK*$_T$4 jets
  - Veto on bad loose
  - Leading jet $P_T > 45$ GeV
  - JVF < 0.75
  - $|\eta| < 2.5$
  - $\Delta R > 0.7$
- Fat jets
  - $|\eta| < 2.5$

Other selections

- Primary vertex
  - At least 3 reconstructed tracks
- Triggers (depending on period)
  - Electrons: EF_e20_medium, EF_e22_medium, EF_e22vh_medium1 || EF_e45_medium1
  - Muons: EF_mu18_MG, EF_mu18_MG_medium