$ZH \rightarrow \nu\nu bb$ in ATLAS

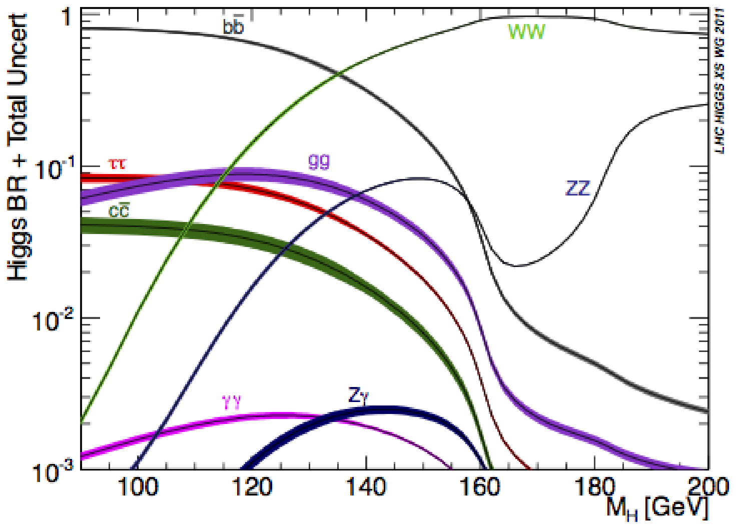
Matthew Jackson
University of Liverpool
IoP – April 2014

Outline

- Brief overview of the ZH → υυbb (0-Lepton) ATLAS analysis
  - Motivation
  - Analysis strategy
  - Status and results
- Monte Carlo based *b*-tagging efficiency studies
  - Motivation
  - Method
  - Overview of results

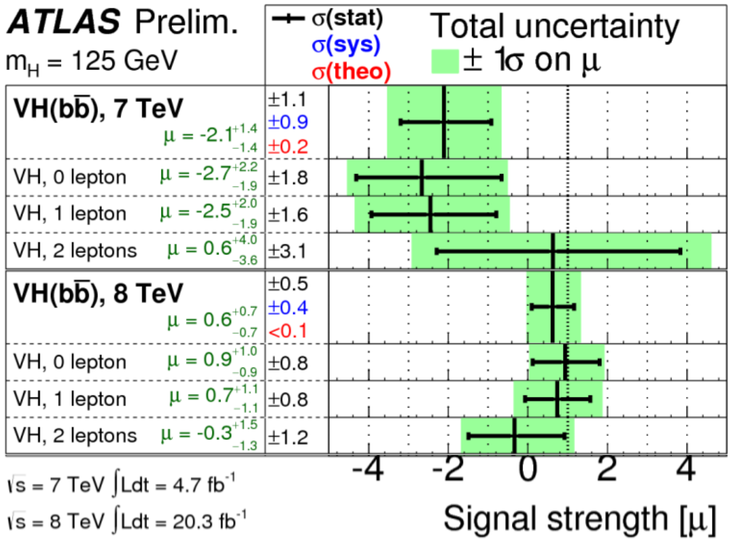
H→*bb* Motivation

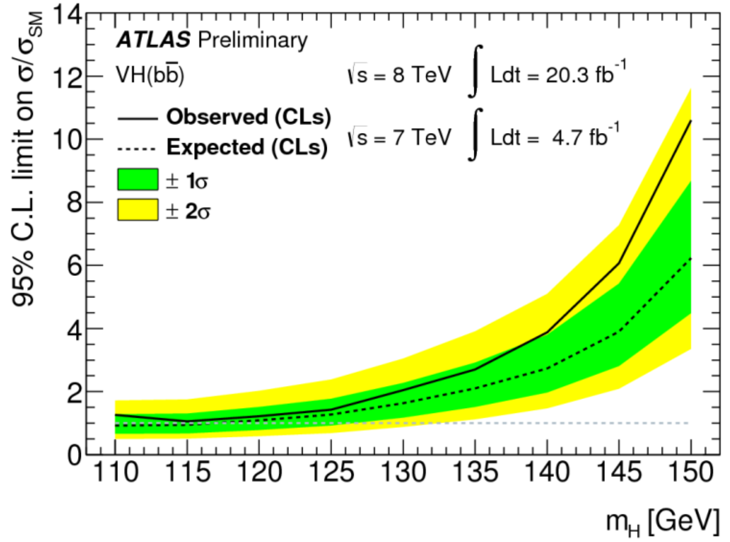
- H → *bb* dominates at low masses
- High multi-jet background makes direct analysis impossible
- Associated production mechanisms are utilised
  - *ZH*, *WH* and *ttH*
- Can supply a direct constraint on coupling to fermions
  - Measure *b* coupling
  - Input to measuring *VH* and *ttH* couplings



Search Strategy

- Three channels are simultaneously explored and combined to increase sensitivity of the VH(*bb*) analysis (0-, 1- and 2- lepton)
  - $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow l\nu b\bar{b}$ and $ZH \rightarrow llb\bar{b}$, my talk focuses on the 0-lepton analysis
  - Work ongoing on cut based and MVA analysis
    - Cut based: Re-optimise EPS 13 analysis
    - Improved with boosted decision tree
  - Many major updates looking for increased sensitivity from EPS (right)

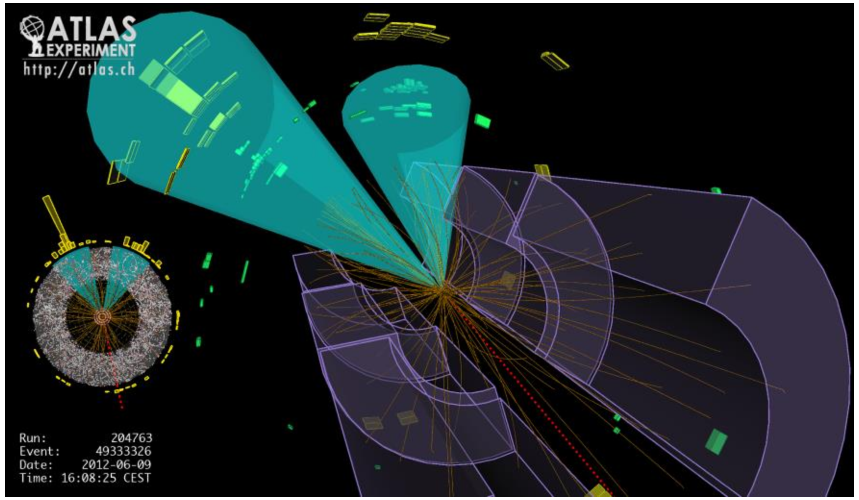


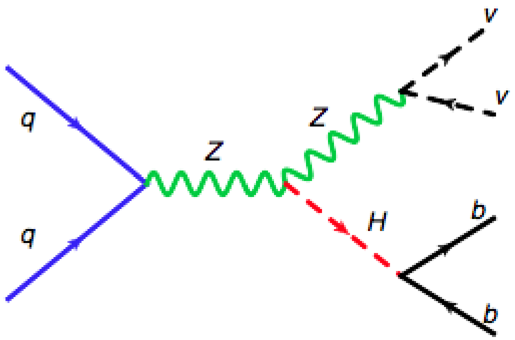


ATLAS-CONF-2013-079
http://cds.cern.ch/record/1563235

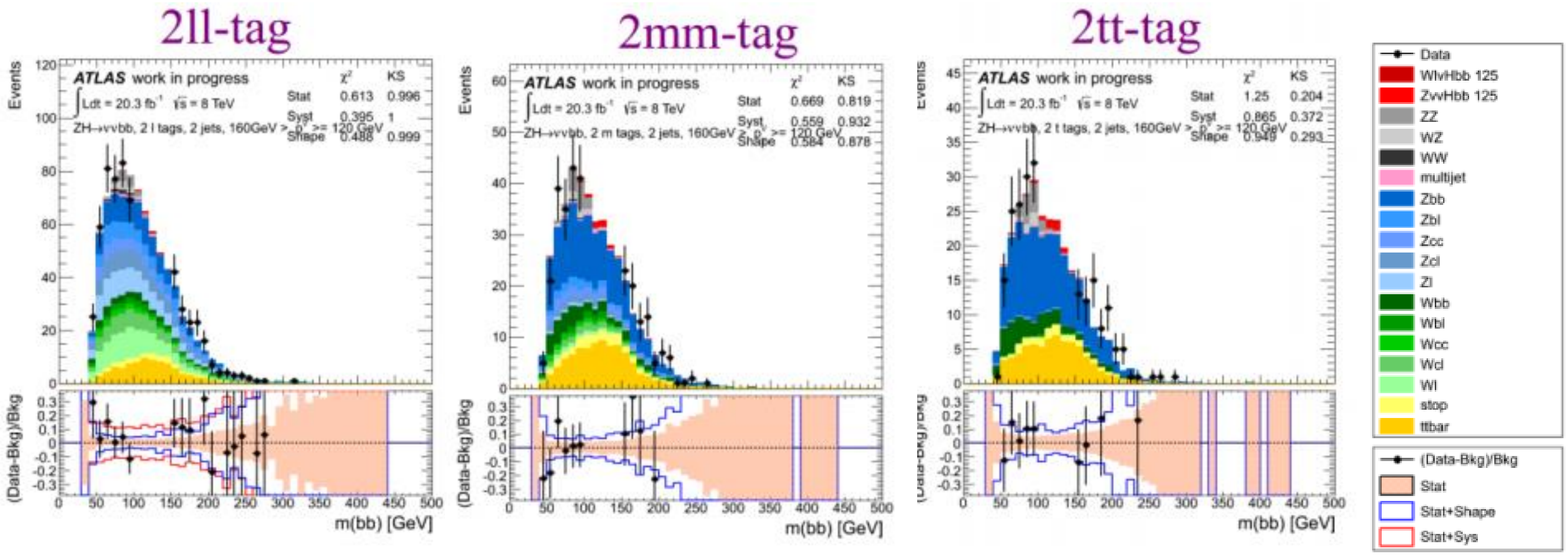
0-Lepton Selection

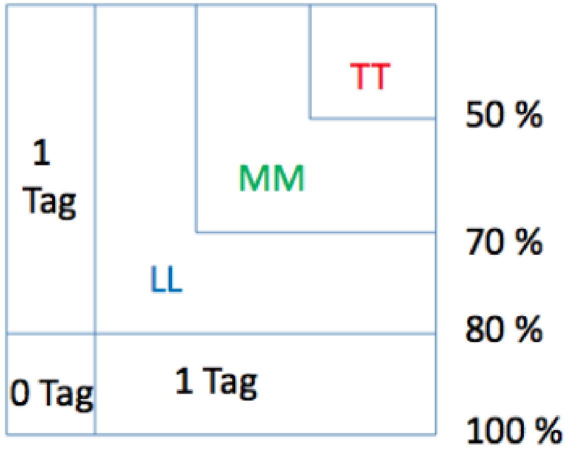
- Veto on loose leptons
- Exactly 2-signal jets tagged with MV1c (multivariate *b*-tagger)
  - Split into 3 categories
  - Two tight (tt), medium (mm) and loose (ll) tagged (50%, 70% and 80% working points)
- E_T^{Miss} > 90 GeV
  - Split in to categories to increase sensitivity
  - Calorimeter E_T^{Miss}
- p_T^{Miss} > 30 GeV
  - Track based E_T^{Miss}
- Topological cuts
  - Reduce QCD backgrounds
  - Exploit different signal and background kinematics
- QCD background derived from data driven method





Results





- Fit using m_{jj} (cut based) or MVA output (MVA) distribution
  - Above plots show 0-lepton, 2-jet, 2-tagged categories
  - Data blinded in sensitive region
- Analysis performed simultaneously in 0-, 1- and 2- lepton channels
  - Additional categories are used to determine backgrounds

Expected Sensitivity

- Many improvements made since EPS analysis
  - MVA analysis
  - Object selection
  - Background modelling
  - Continuous b-tagging
  - Understanding of systematics
  - Many more..
- Expected sensitivity can be estimated using post fit plots

$$\sum_{bins} \frac{Signal}{\sqrt{Background}}$$

	EPS	Cut Based	MVA
0-lepton	1.33	1.54	1.81
1-lepton	1.26	1.50	1.93
2-lepton	0.88	0.99	1.32
Combined	2.03	2.36	2.96

b-tagging Efficiency – MC Studies

- *b*-tagging in many analyses, including VH, is a dominant systematic
  - Particularly when high pT jets are involved
- Method
  - Calculate tagging efficiency for selected tagging algorithm as a function of pT using tt MC
  - Consider sources of uncertainty
  - Calculate total uncertainty in each pT bin
- Using MC to study b-tagging efficiency it is possible to extend the pT range available to data driven studies
  - Reduce systematic uncertainty
  - Improve statistical uncertainty

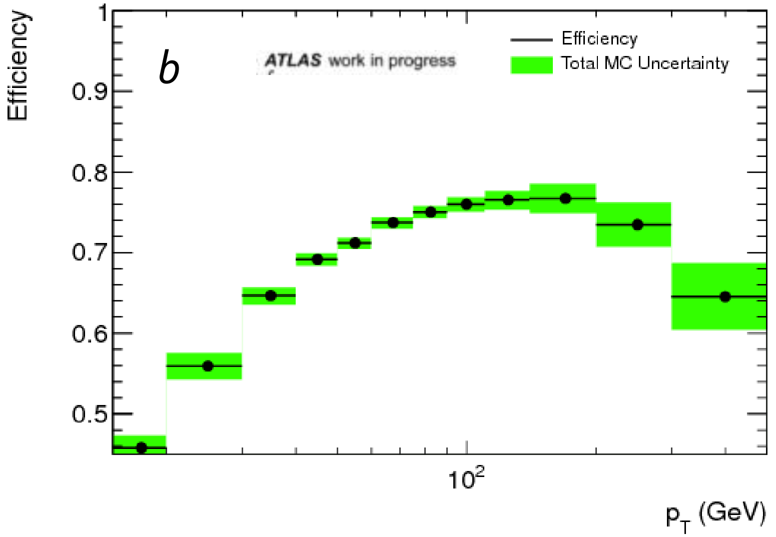
Uncertainty [%]	0 lepton
b-tagging	2.96
c-tagging	2.17
light tagging	1.21
Jet/Pile-up/E_T^{miss}	2.36
Lepton	0.42
Top modelling	1.18
W modelling	1.75
Z modelling	2.67
Single-top modelling	0.20
Diboson	0.49
Multijet	0.00
Luminosity	2.32
Total	2.36

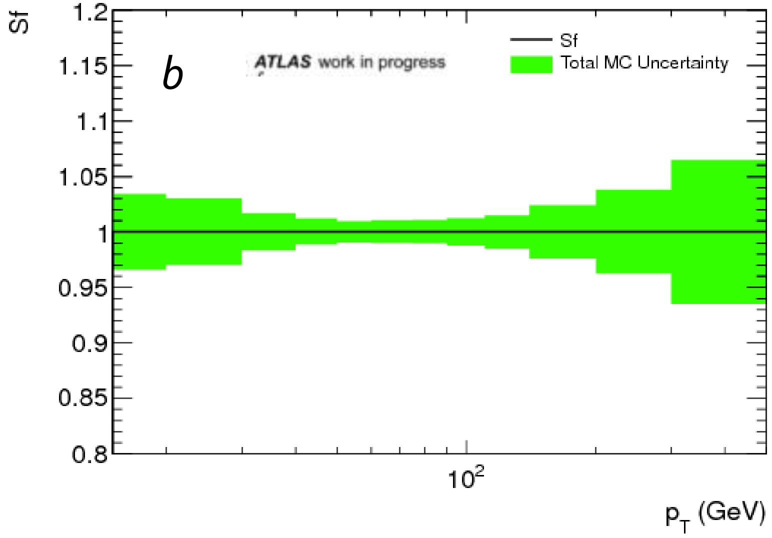
Uncertainty [%]	0 lepton
b-tagging	5.58
c-tagging	0.02
light tagging	0.00
Jet/Pile-up/E_T^{miss}	5.43
Lepton	0.00
VH p_T–dependence	2.03
VH theory scale	1.35
VH theory PDF	3.01
Luminosity	2.80
Total	9.39

ATLAS-CONF-2013-079
http://cds.cern.ch/record/1563235

Results
b

Results are shown for the 2012 analysis using the MV1c tagger at 70% efficiency
Uncertainties stated are relative to nominal value

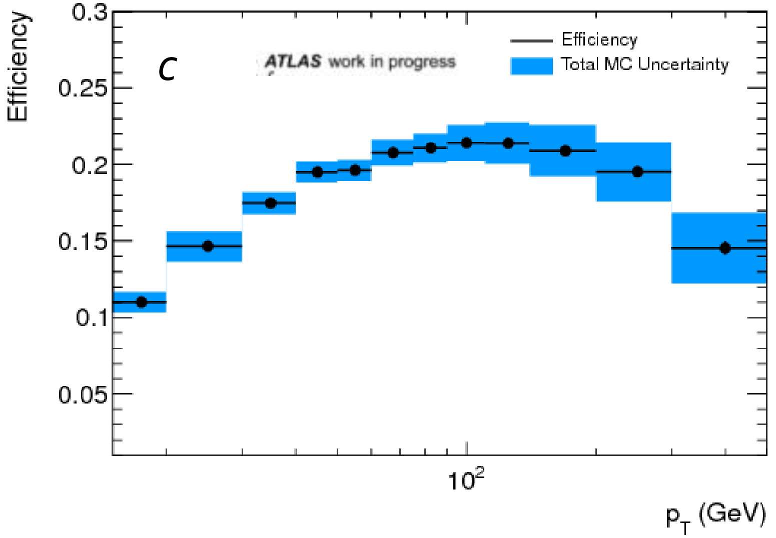


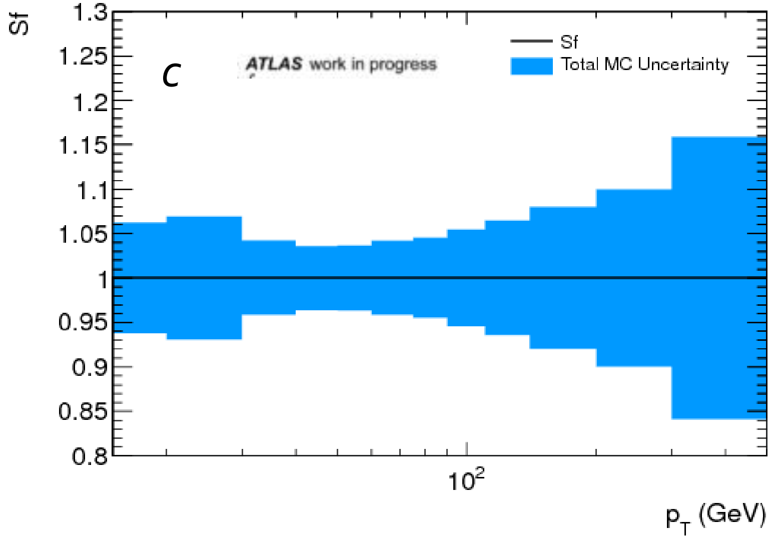


- Range from 1.0% (50 – 60 GeV) to 6.5% (300 – 500 GeV)
  - Jet energy systematics dominate first three p_T bins
  - Shared hits and fake tracks contribute more for p_T > 50 GeV

Results
c

Results are shown for the 2012 analysis using the MV1c tagger at 70% efficiency
Uncertainties stated are relative to nominal value

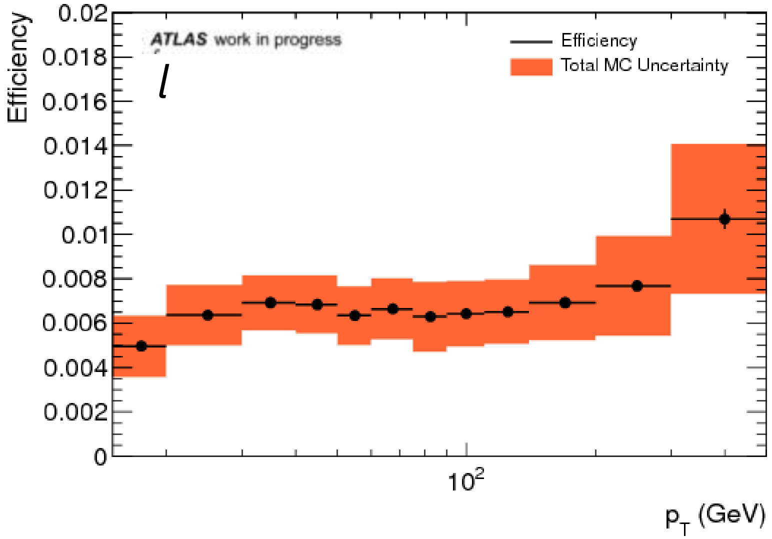


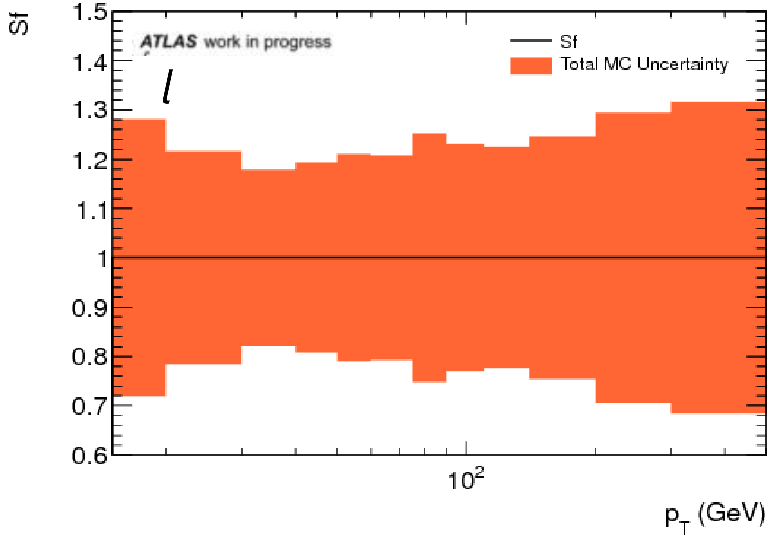


- Range from 3.6% (40 – 50 GeV) to 15.9% (300 – 500 GeV)
  - Similar story to *b*
  - Shared hits and fake tracks dominate at high p_T

Results
|

Results are shown for the 2012 analysis using the MV1c tagger at 70% efficiency
Uncertainties stated are relative to nominal value

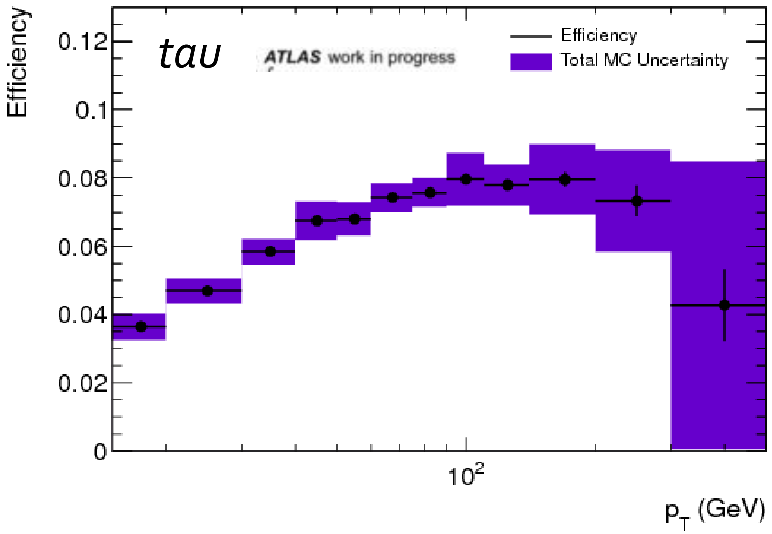


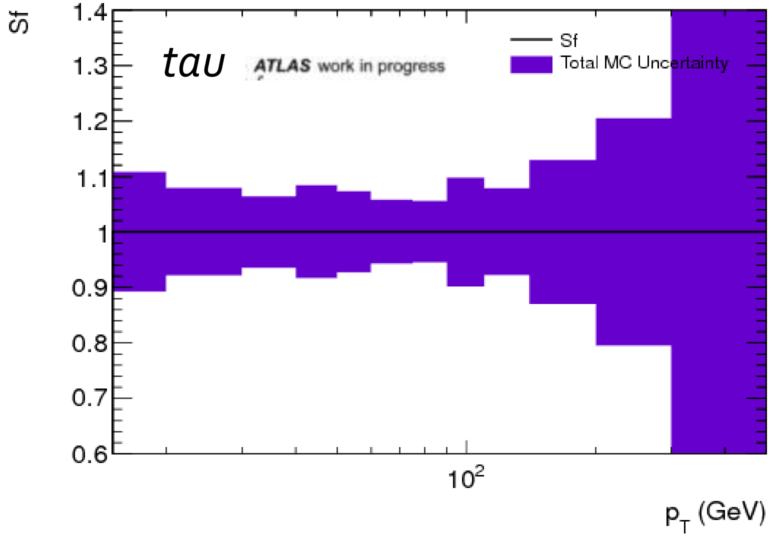


- Range from 17.9% (30 – 40 GeV) to 31.7% (300 – 500 GeV)
  - Jet energy systematics still contribute significantly
  - Fake tracks dominates across the p_T range

Results
tau

Results are shown for the 2012 analysis using the MV1c tagger at 70% efficiency
Uncertainties stated are relative to nominal value





- Range from 5.6% (60 – 75 GeV) to 20.5% (200 – 300 GeV)
  - Low stats, particularly at high p_T
  - Similar story to light

Summary

- MC based *b*-tagging efficiency studies can be used to reduce systematic uncertainties
  - Improve statistics
  - Extend p_T range of data driven methods

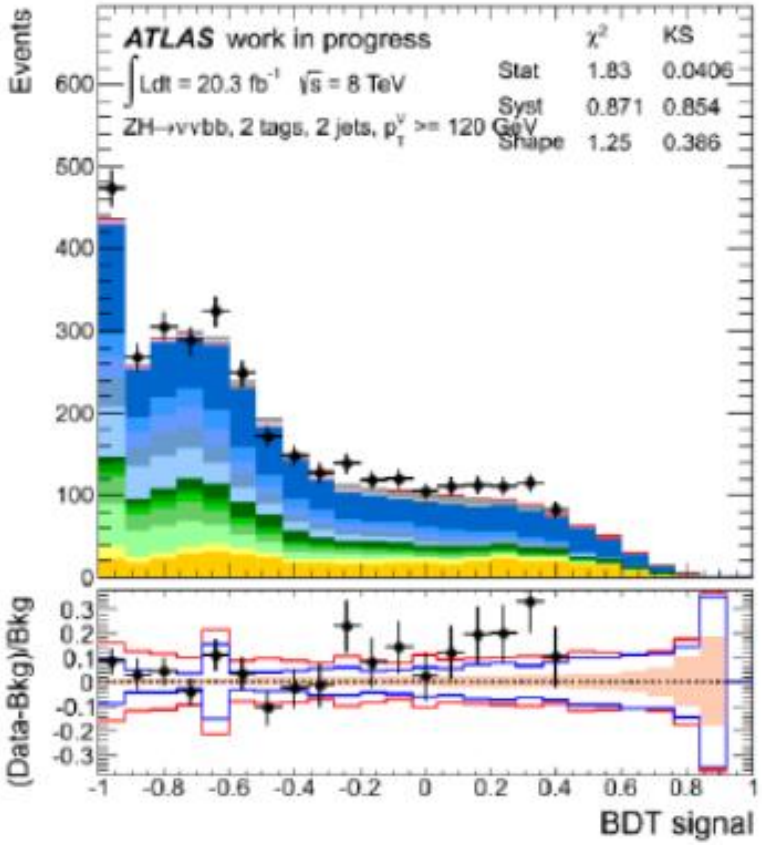
- Currently being incorporated into ATLAS calibration data interface (CDI) for use in any analysis
  - Internal documentation

Back up

MVA Inputs

Variable		0-lepton	1-lepton	2-lepton
$\Delta R(bb)$				
$abs(\Delta\eta(bb))$				
$abs(\Delta\eta(V,bb))$				
$abs(\Delta\phi(V,bb))$				
$abs(\Delta\phi(l,\text{closest } b))$				
p_T^V				
M_{bb}				
$HT(l+E_T^{Miss}+jets)$				
p_T^{b1}				
p_T^{b2}				
m_T^W				
E_T^{Miss}				
m_{ll}				
$MV1c_{b1}$				
$MV1c_{b1}$				
3-jet	m_{bbj}			
	p_T^{b3}			

MVA Output



ATLAS work in progress
∫Ldt = 20.3 fb⁻¹ √s = 8 TeV
ZH→ννbb, 2 tags, 2 jets, p_T^V >= 120 GeV
χ² KS
Stat 1.83 0.0406
Syst 0.871 0.854
Shape 1.25 0.386
Events
(Data-Bkg)/Bkg
BDT signal

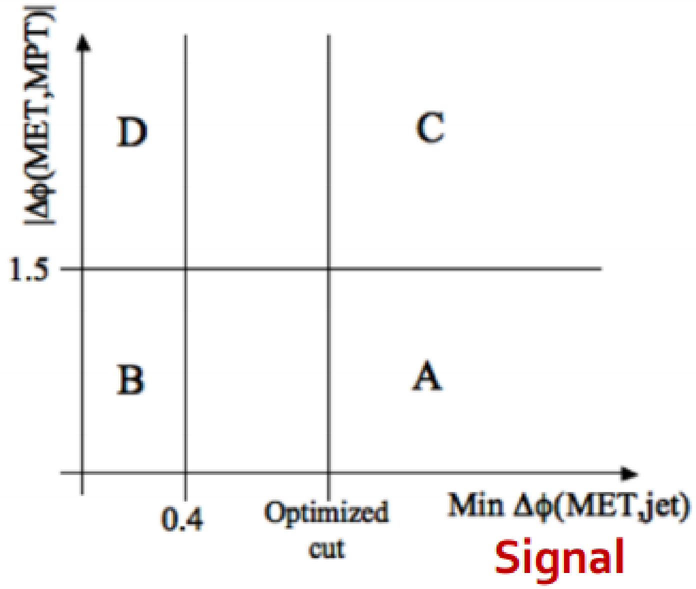
Global Fit

- Perform analysis simultaneously in 0-, 1-, and 2-lepton channels
  - Additional categories are used to determine backgrounds; 2 and 3 jets; 1 and 2 b-tagged jets
  - Fit mjj or mva distribution
- Cut based
  - Fit LL/MM/TT in 5 pTV bins (0-90, 90-120, 120-160, 160-200, 200+)
    - 0-lepton starts at 100 GeV (pTV ~ MET)

	Control		Signal		Control	
	2 Jet, 1 b-Tag	3 Jet, 1 b-Tag	2 Jet, 2 b-Tag	3 Jet, 2 b-Tag	Top e-μ CR	
0-lepton	W+jets/Z+jets	W+jets/Z+jets/Top	Z+b/Top	Z+b/Top	-	3 E_T^{Miss} bins
1-lepton	W+c	W+c / Top	W+b / Top	Top	-	5 p_T^V bins
2-lepton	Z+c	Z+c	Z+b	Z+b	Top	5 p_T^V bins

Multi-jet Background Estimation

- **0 Lepton**
  - ABCD method
  - Regions defined by relative directions of MET/jets/p_T^{miss}
  - Small (~1%)
- **2 Lepton**
  - Template: Reverse lepton isolation and loosen quality selection
    - Fit to m_{ll} sidebands
  - Negligible (< 1%)
- **1-lepton**
  - Template: Revers track isolation
  - Remove HT and MET cuts
    - Fit shape of pT(lep)



$$N_{QCD}(A) = \frac{N(B)}{N(D)} \times N(C)$$

b-tagging Efficiency – Considered Uncertainties [1]

- Systematic uncertainties considered in 2012 analysis
  - IP Smearing
    - Smear core of d_0 and z_0 distributions by values obtained from the insitu-database
    - Work on long range smearing study on going through other studies
  - Fake tracks
    - Many sources e.g. material interactions, dead sensors
    - Throw away a random 50% of tracks (track probability < 0.8)
  - Shared hits
    - 1 or more shared hits in the pixel detector
    - 2 or more shared hits in SCT
    - Throw away a random 50% of tracks with shared hits
  - Track multiplicity (measured in *b*-jet)
    - Re-weight by scale factor obtained in ZH analysis track multiplicity distribution in a top dominated region
    - Accounts for uncertainty in track efficiency and the number of stable charged particles coming from the *b*

b-tagging Efficiency – Considered Uncertainties [2]

- Systematic uncertainties considered in 2012 analysis
  - Jet systematics
    - Jet axis resolution
      - Smear η and ϕ with a Gaussian
    - Jet energy systematics
      - Full treatment of 51 shifts as prescribed by the ATLAS jet group
  - *b*-fragmentation modelling
    - Jet-by-jet reweight based on comparison to experimental values of the *b*-fragmentation function
    - Not yet included in 2012 results
  - Pile up
    - Re-weight by pile up weight
    - New for 2012 analysis