Background modelling and evaluation of uncertainties in VH(bb) searches

Inês Ochoa

16th April 2015

First Annual Meeting of ITN HiggsTools, University of Freiburg

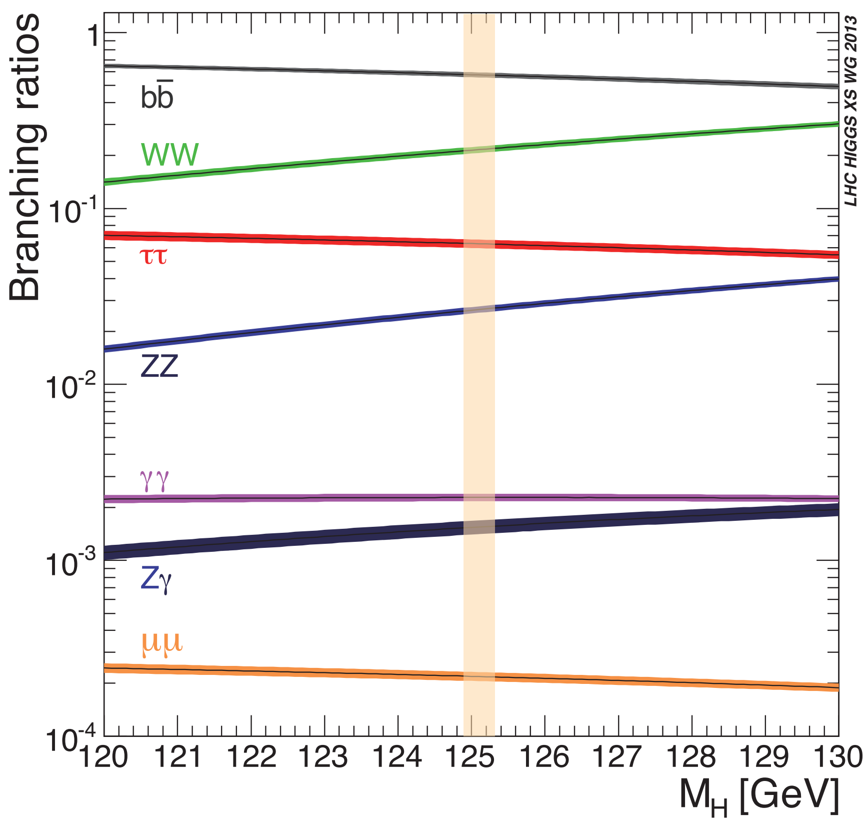
Outline

- H→bb: motivation
- Summary of ATLAS and CMS VH(bb) analyses
  - Strategy and results
- Overview of background modelling in ATLAS and CMS
  - Corrections to simulation
  - Systematic uncertainties
- Summary and conclusions

Searching for H→bb

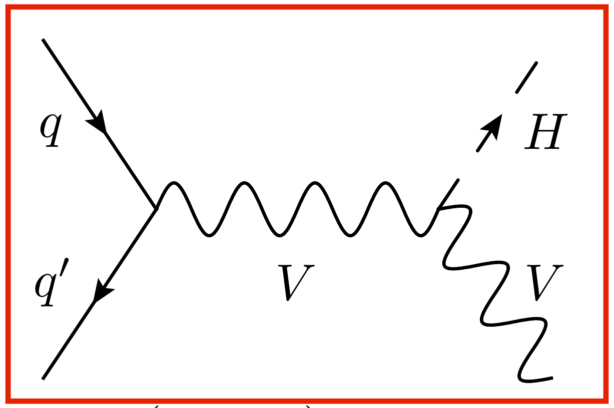
Why?

- The discovered Higgs boson is so far consistent with a Standard Model Higgs
- One crucial piece missing: evidence of coupling to bottom-quarks
- H→bb is the decay channel with the largest branching ratio (~60%)
- Observation and measurement of this decay is crucial for constraining coupling measurements and determining the nature of the Higgs.



How?

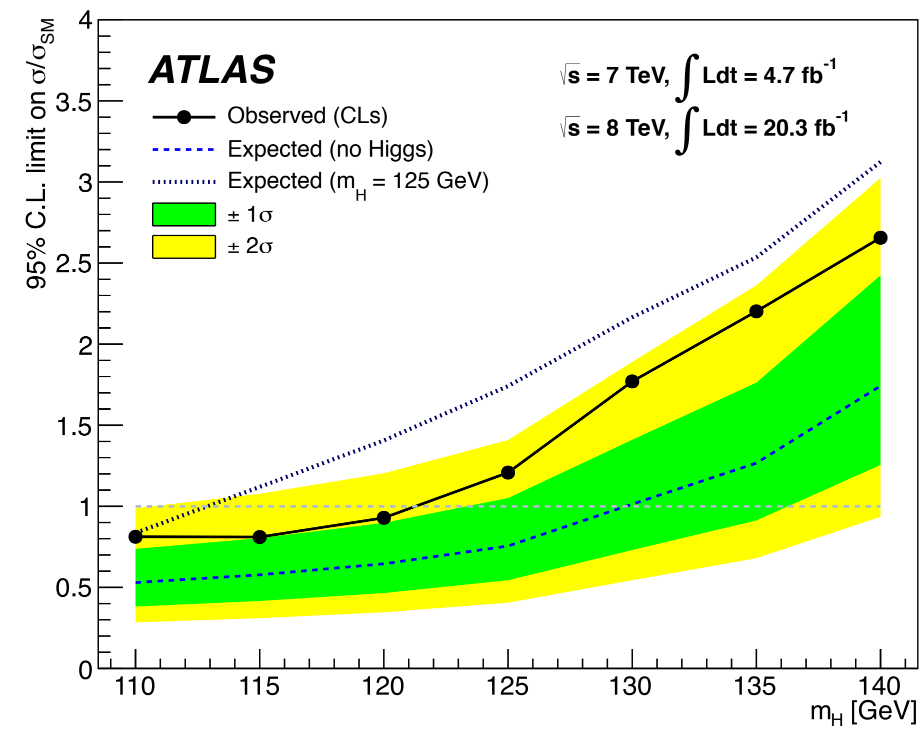
- Despite largest branching ratio, there are plenty of jets in a hadron collider...
- Need a handle on QCD multi-jet backgrounds and something to trigger on: look for associated production with a vector boson



$\sigma_{VH}(8 \text{ TeV}) \approx 1.12 \text{ pb}$

ATLAS VH(bb) in a nutshell

- VH(bb) analysis of Run 1 dataset published in JHEP: JHEP 01(2015) 69

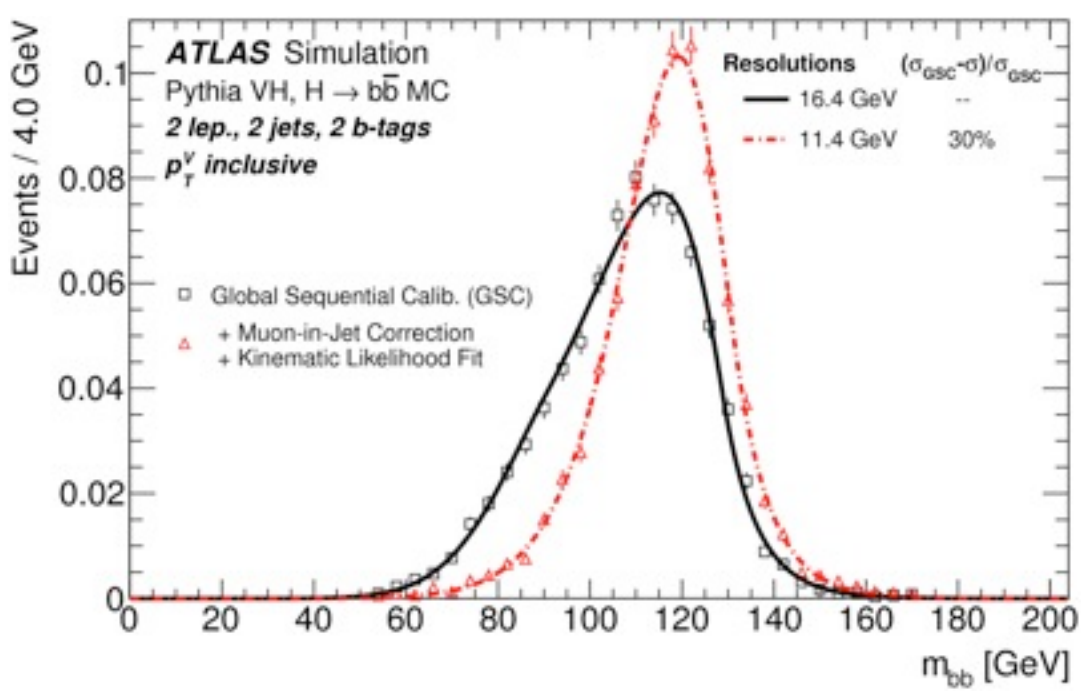


- 8 TeV data: based on Boosted Decision Tree (BDT)
- 7 TeV data: based on the dijet invariant mass (m_{bb})

- Small observed (expected) excess of 1.4σ (2.6σ)

- Signal strength: $\mu = 0.51 \pm 0.31(\text{stat.}) \pm 0.25(\text{syst.})$

$$\mu = \sigma/\sigma_{SM}$$

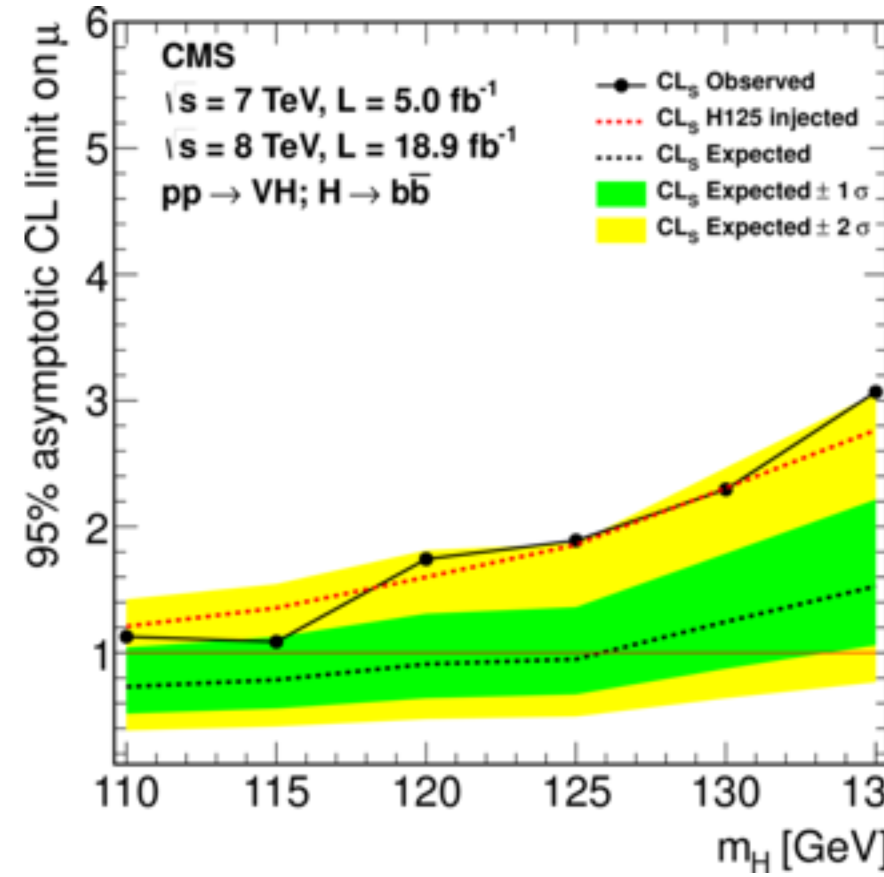


- BDT trained against sum of backgrounds, on dijet mass, jet and lepton kinematic variables and b-tagging discriminants

CMS VH(bb) in a nutshell

- VH(bb) analysis of Run I dataset published in: PRD 89 (2014) 012003

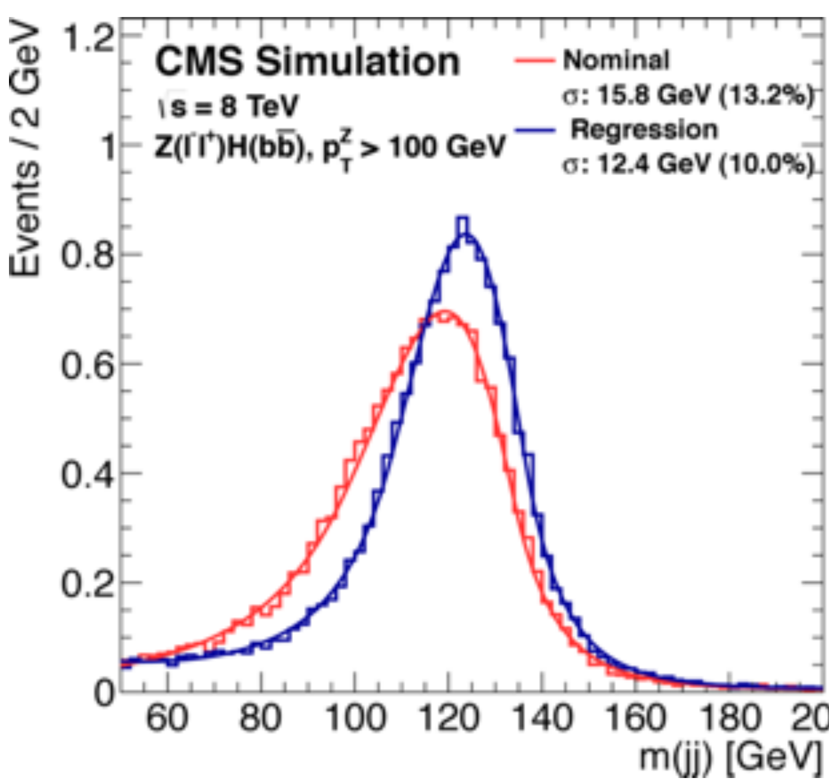
- Analysis of 7 and 8 TeV datasets with multivariate analysis
  - Specific BDTs trained against specific background processes (V+jets, top, diboson)
  - Results extracted from fit to final BDT distribution



- Small observed (expected) excess of 2.0σ (2.5σ)
- Signal strength: $\mu = 0.83 \pm 0.35$

$$\mu = \sigma/\sigma_{SM}$$

(including contribution from ttH(bb) and gg→ZH)



- Both experiments use techniques to improve b-jet energy calibration and improve dijet mass resolution

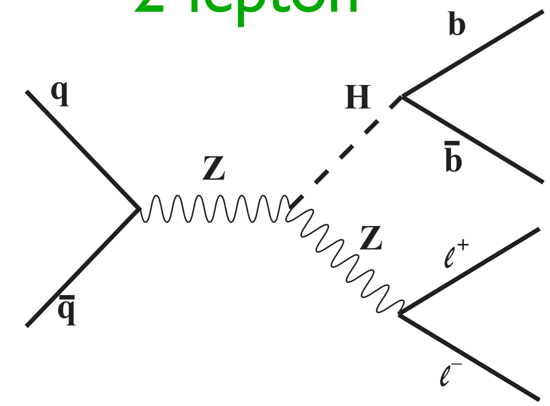
Final state topologies

- ATLAS: 3 lepton selections, targeting Z→νν, W→lν, Z→ll
- CMS: W(μν), W(eν), W(τν), Z(μμ), Z(ee) and Z(νν)
- Two jets originating from b-quark fragmentation

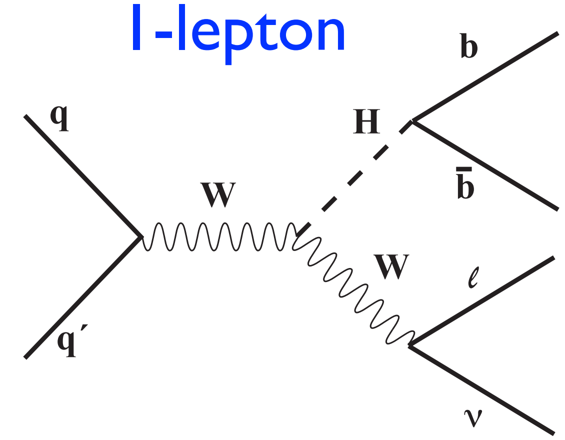
- Several SM processes can mimic these channels:
  - Top-pair and single-top
  - W + bb, Z + bb (irreducible)
  - Diboson (WW, WZ, ZZ)

- Experimental backgrounds:
  - W+jets and Z+jets where a light/charm-jet fakes a b-jet
  - Multijet (fake leptons, semileptonic heavy-flavour decays, fake missing transverse energy)

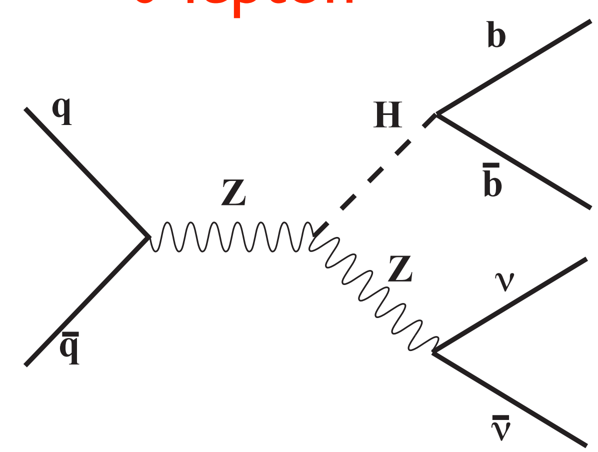
2-lepton



1-lepton



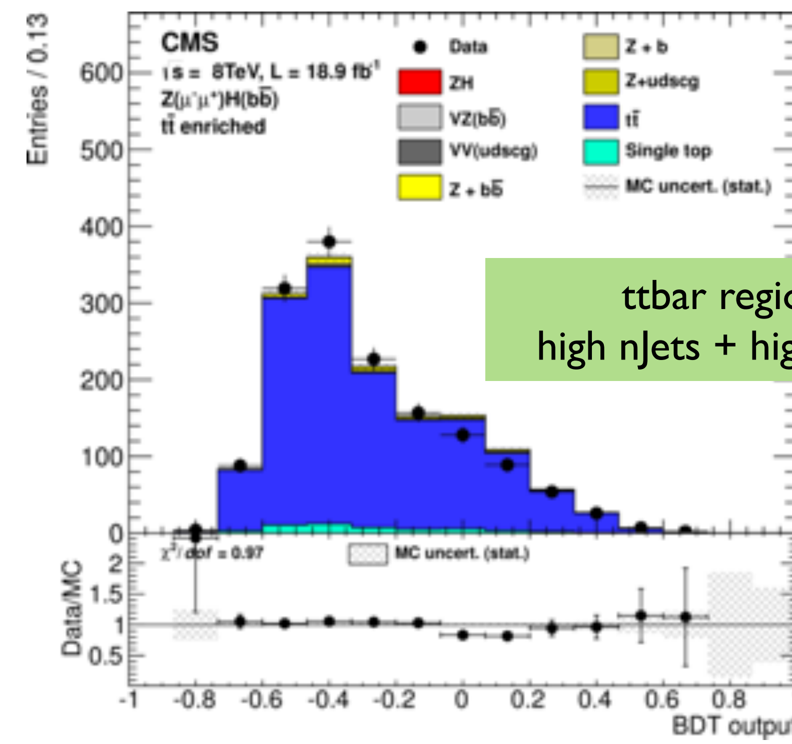
0-lepton



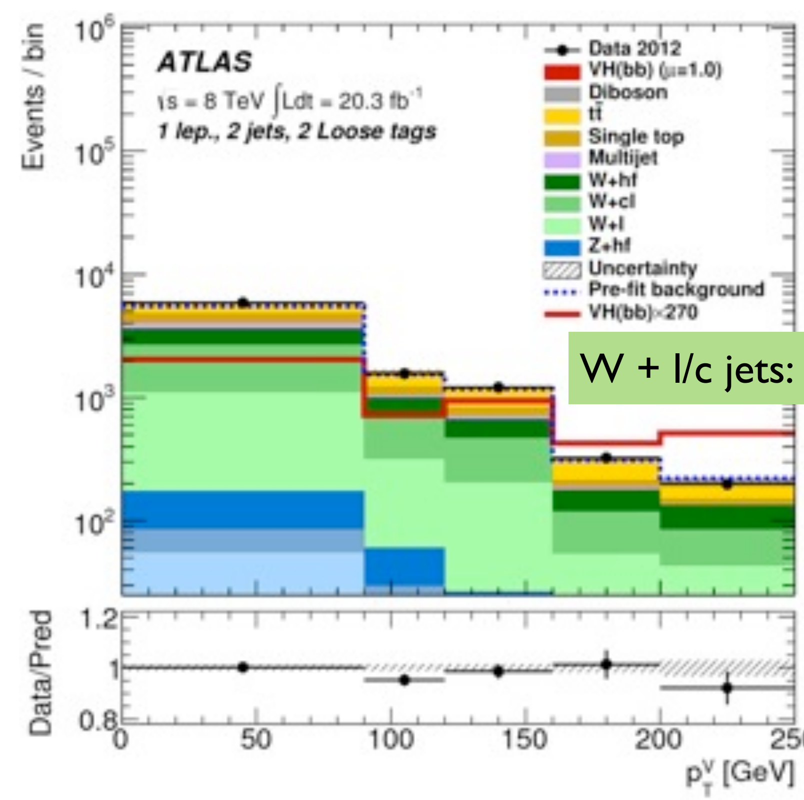
Many different, large backgrounds, probed over a wide range of phase-space.

Analysis strategy

- The signal has a harder pT(V) spectrum than the background:
  - take advantage of higher S/B in **boosted** region
  - categorize in pT(V) bins
- To be able to get a handle on the different background processes, constrain normalizations and shapes:
  - define control regions or categorize in other variables
- Main variables of interest:
  - jet multiplicity
  - number of b-tagged jets and b-tagging operating points (purity)
  - pT(V) distribution
  - m_{bb} distribution



ttbar region:
high nJets + high b-tag

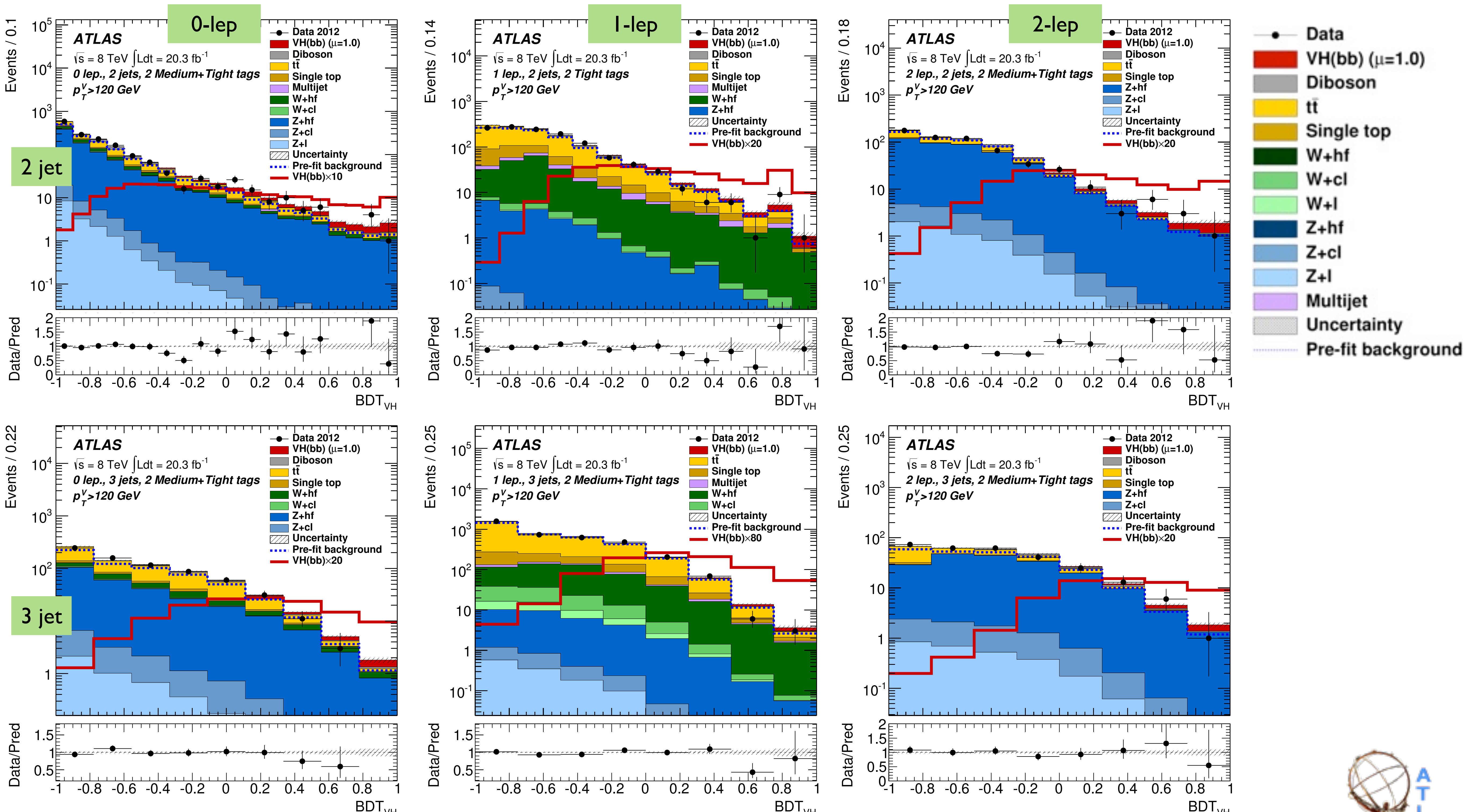


W + l/c jets: low b-tag

Analysis strategy

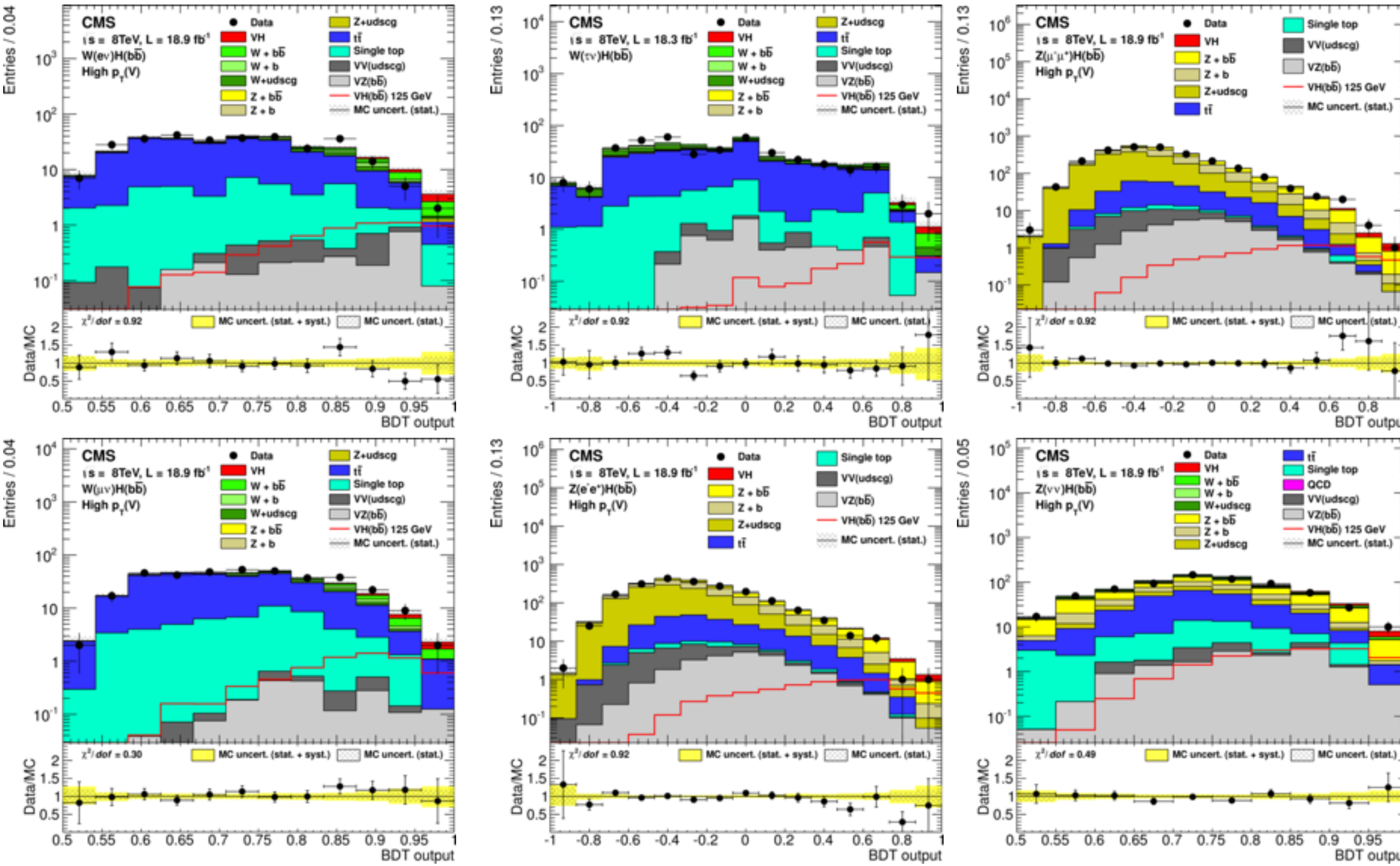
38 event regions*

*only most sensitive regions shown here: high boost + high b-tagging purity



14 event regions*

*only most sensitive regions shown here: high boost + vh-enriched region



Background uncertainties

- Analysis with low signal-to-background ratio
  - Background uncertainties are a key component
- How can each background impact the final result?
  - Migration between event categories
  - Shifts on BDT/m_{bb} shape

Main variables: m_{bb}, pT(V), 3-to-2 jet ratio

- Strategy: parameterize main discrepancies found in data-to-MC and MC-to-MC comparisons

ATLAS

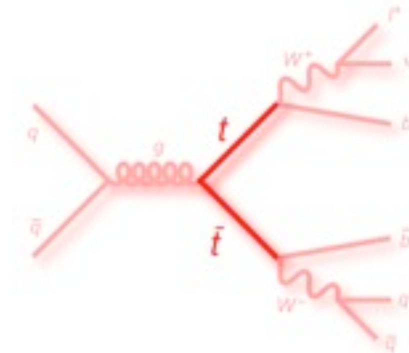
Process	Generator
Signal$^{(\star)}$	
$q\bar{q} \to ZH \to \nu\nu bb/\ell\ell bb$	PYTHIA8
$gg \to ZH \to \nu\nu bb/\ell\ell bb$	POWHEG+PYTHIA8
$q\bar{q} \to WH \to \ell\nu bb$	PYTHIA8
Vector boson + jets	
$W \to \ell\nu$	SHERPA 1.4.1
$Z/\gamma* \to \ell\ell$	SHERPA 1.4.1
$Z \to \nu\nu$	SHERPA 1.4.1
Top-quark	
$t\bar{t}$	POWHEG+PYTHIA
t-channel	ACERMC+PYTHIA
s-channel	POWHEG+PYTHIA
Wt	POWHEG+PYTHIA
Diboson$^{(\star)}$	POWHEG+PYTHIA8
WW	POWHEG+PYTHIA8
WZ	POWHEG+PYTHIA8
ZZ	POWHEG+PYTHIA8

CMS

Process	Signal	top-pair	V+jets	diboson	single-top
Generator	Powheg +Herwig++	MadGraph* +Pythia	MadGraph* +Pythia	MadGraph* +Pythia	Powheg +Herwig++

*version 5.1

ATLAS: top-pair production

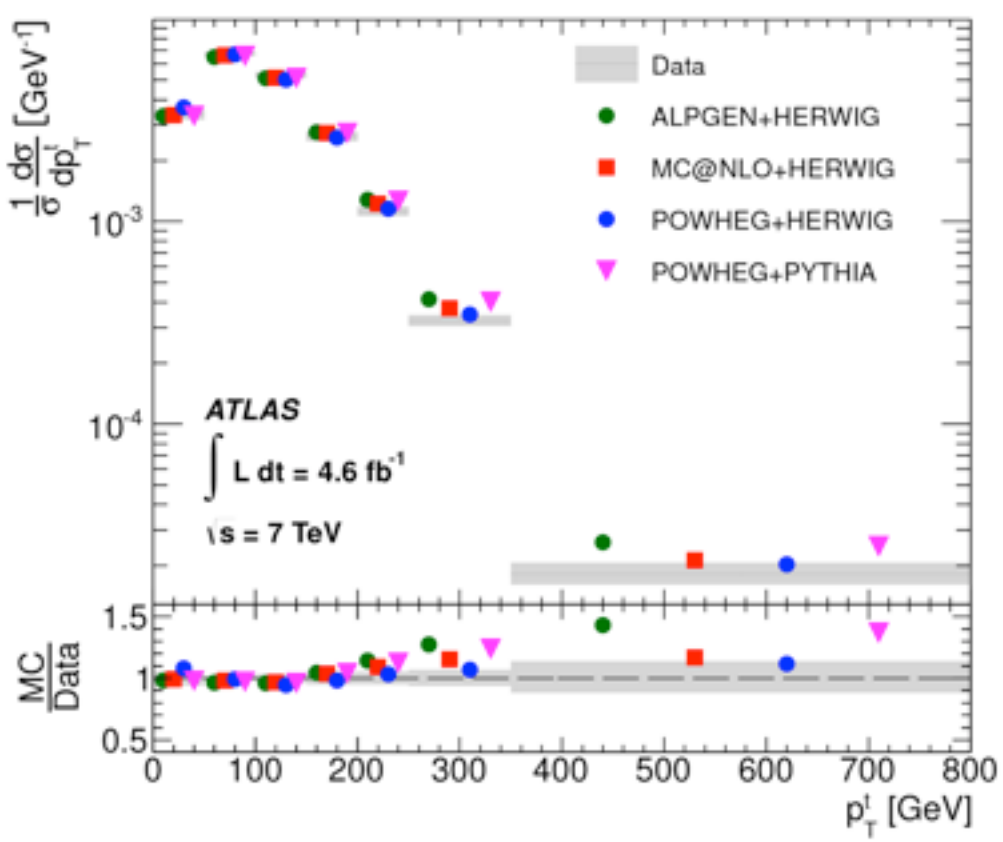


Introduction

- Powheg + Pythia as baseline generator
- CT10 PDFs for matrix element, CTEQ6L1 with Perugia2011C for parton shower
- Cross-section @ NNLO with NNLL resummations used for selection optimisation
- Each lepton channel accesses very different final-states of top-pair phase-space:
  - 0- and 1-lepton channels: most likely jet from W decay is missed
  - 2-lepton channel: likely that an ISR/FSR jet is selected

Correction to simulation

- Reweighting of pT(top) at generator-level, based on ATLAS measurement
- Assigned systematic of 50% of correction



ATLAS: top-pair production

Shape uncertainties on main variables

- Generator comparisons based on 1-lepton channel:
  - PDF variations: nominal vs HERAPDF
  - Parton shower and hadronisation scheme: nominal vs Herwig
  - NLO ME and matching scheme: nominal vs MC@NLO+Herwig
  - ISR/FSR: AcerMC + Pythia with different rates
  - higher order tree-level ME: nominal vs Alpgen + Pythia

In general, largest deviations observed for Alpgen.

- e.g. m_{bb} shape: ±3% variations @ 50 GeV and ∓1% @ 200 GeV (for 2/3-jet regions)

$t\bar{t}$	
3/2-jet ratio	20%
High/low-p_T^V ratio	7.5%
Top-quark p_T, m_{bb}, E_T^{miss}	S

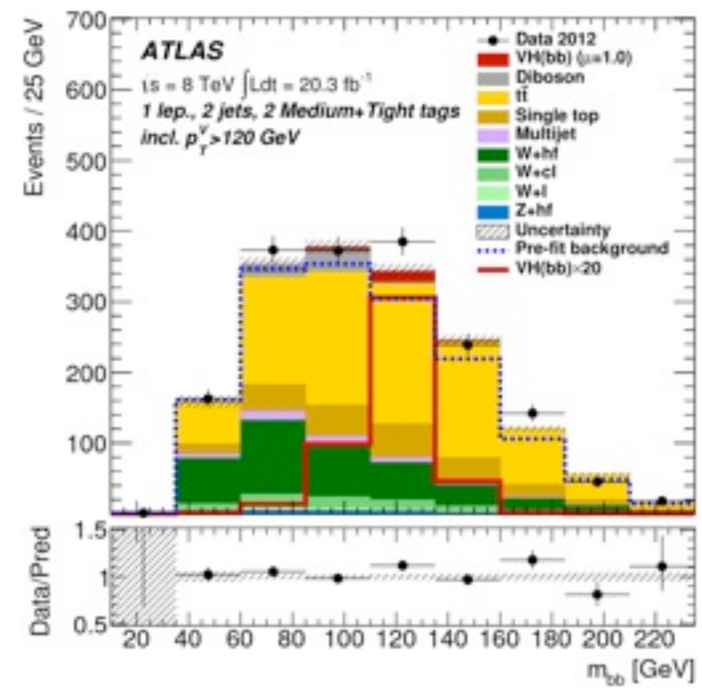
ATLAS: top-pair production

Normalization uncertainties

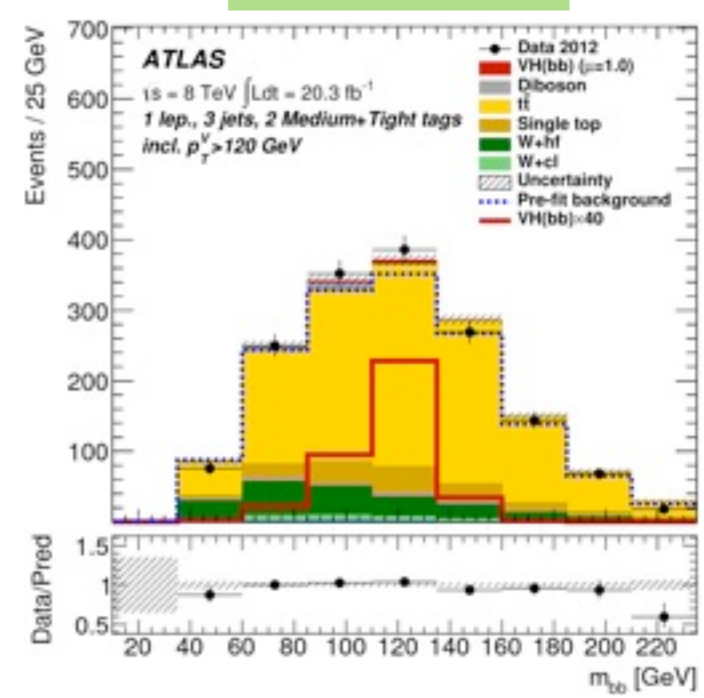
- Freely-floating normalization in the 2-jet region
- 3-to-2 jet uncertainties:
  - 20%, from previously shown generator comparisons
- Independent in each lepton channel
  - different top-pair topologies
  - 3-jet regions very rich in top in 0 and 1-lepton channels
  - m_{ll} sidebands and low BDT values constrain top in 2-lepton channel

Process	Scale factor
$t\bar{t}$ 0-lepton	1.36 ± 0.14
$t\bar{t}$ 1-lepton	1.12 ± 0.09
$t\bar{t}$ 2-lepton	0.99 ± 0.04

1 lepton, 2 jets

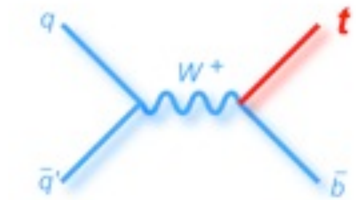


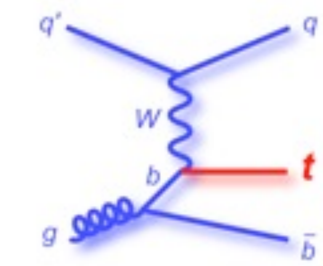
1 lepton, 3 jets



ATLAS: single-top production

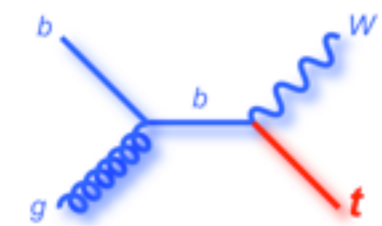
UCL





Introduction

- Powheg+Pythia as baseline generators for s- and Wt-channels
- AcerMC+Pythia as baseline generator for t-channel
- No good control region ⇒ uncertainties from generator comparisons



Normalization and Shape uncertainties

- Cross-section uncertainties from theory

arXiv:1205.3453

	Single top
Cross section	4% (s-,t-channel), 7% (Wt)
Acceptance (generator)	3%–52%
m_{bb}, $p_T^{b_2}$	S

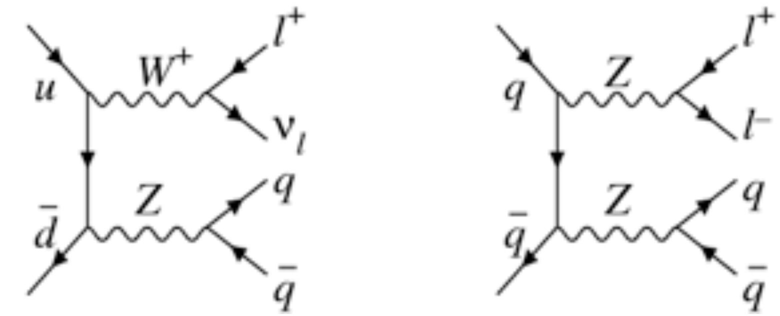
- Acceptance uncertainties on 2/3-jet and low/high pT(V) rates:
  - s-channel: nominal vs MC@NLO vs AcerMC
  - t-channel: nominal vs aMC@NLO+Herwig
  - Wt-channel: nominal vs Powheg+Herwig vs MC@NLO vs AcerMC
    - also Diagram Removal vs Diagram Subtraction (Powheg)
  - for all: ISR/FSR rates with AcerMC samples

t-channel, 2-jet low pT(V)

m_{bb} Wt-channel:
±20% at 50 GeV, ∓40% at 200 GeV

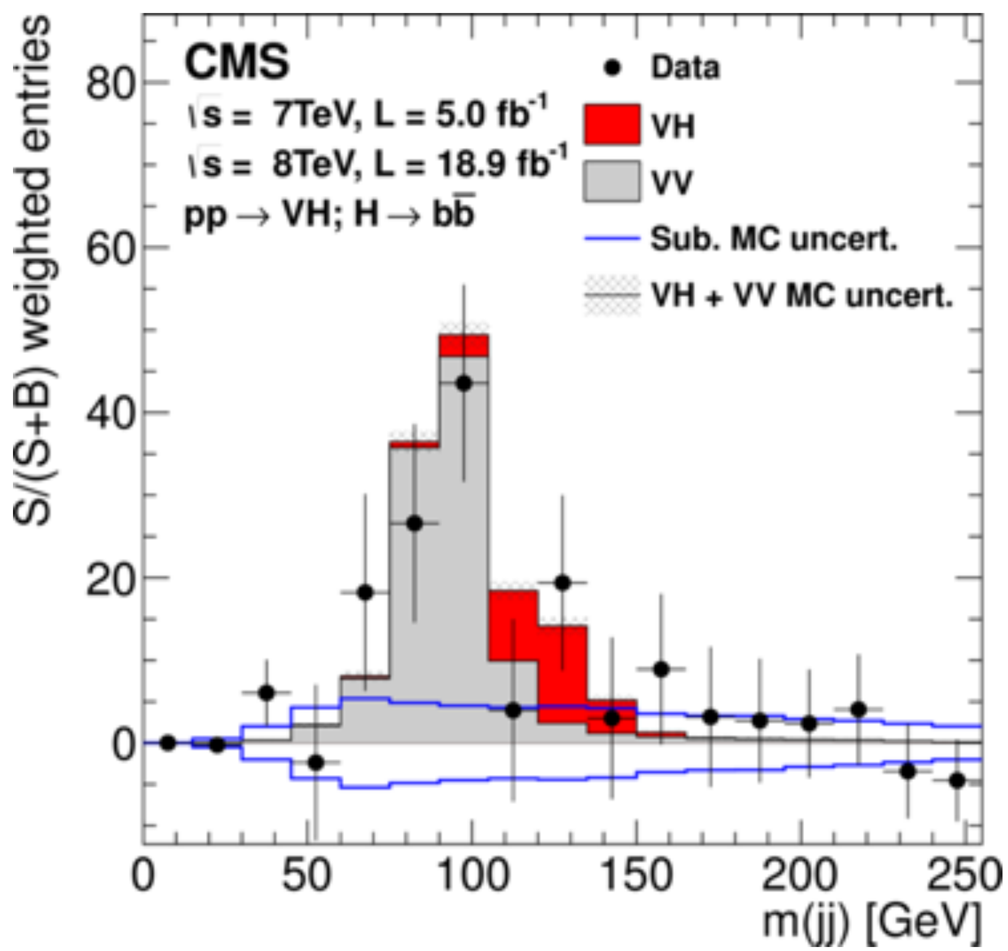
Wt-channel: additional shape uncertainties on m_{bb} and p_T(subleading jet)

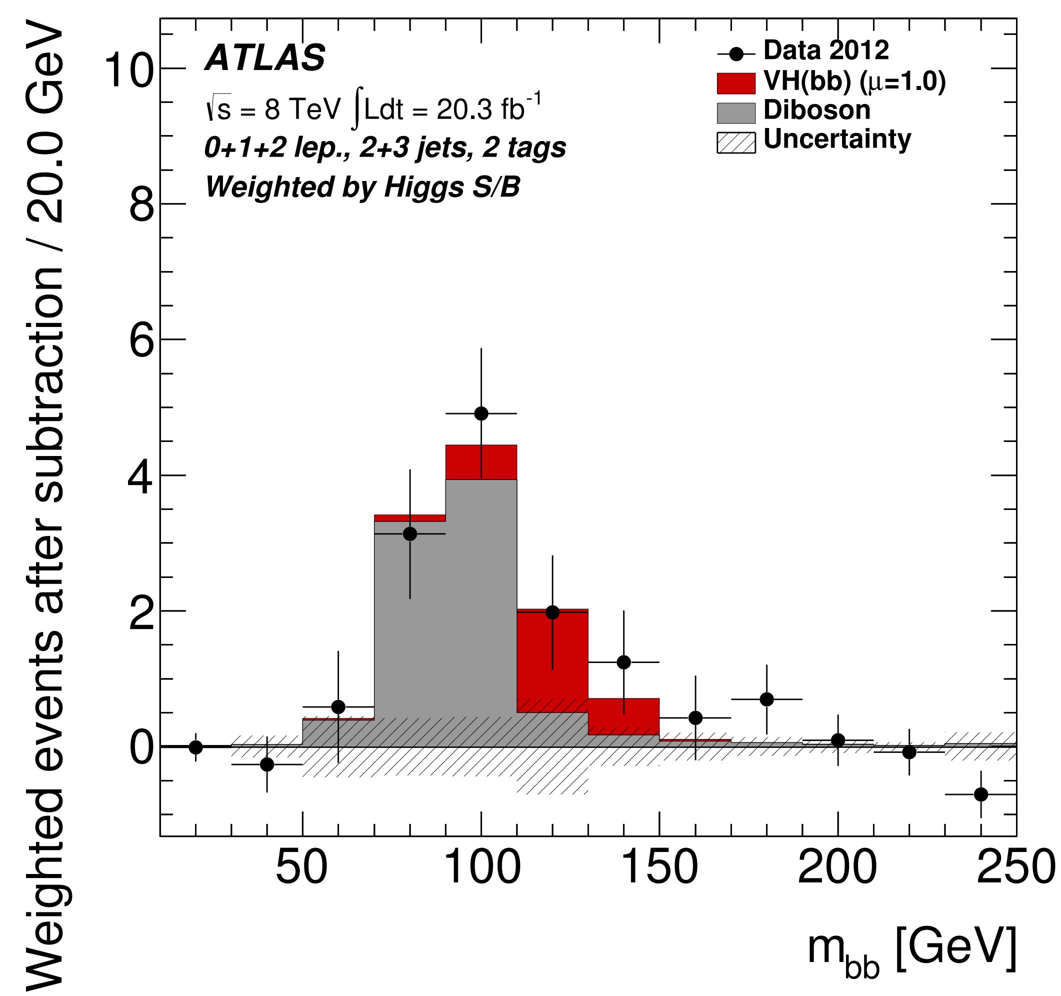
ATLAS/CMS: diboson production



Introduction

- WZ, ZZ: same final state, larger cross-section
- Bonus: train BDT to perform a diboson (WZ,ZZ) measurement
- Validation of the full analysis, including background estimates





Observation of diboson with yields consistent with the SM.

ATLAS: diboson production

Normalization and Shape uncertainties

- Powheg+Pythia as baseline generator (Herwig @ 7 TeV)
- No good control region ⇒ uncertainties from studies at generator level
- Acceptance: parton level with MCFM @ NLO in QCD, in jet and pT(V) bins
  - scale variations μ_R, μ_F by factors 0.5 and 2
  - PDF+α_S variations (PDF4LHC prescription)
  - approximately 20% at high pT(V)
- m_{bb} shape:
  - parton showering/hadronization model
  - nominal vs Herwig
  - 20% effect at 125 GeV

Diboson	
Cross section and acceptance (scale)	3%–29%
Cross section and acceptance (PDF)	2%–4%
m_{bb}	S

ATLAS: W/Z+jets

Introduction

- Sherpa 1.4.1 as baseline generator
- Regions with 0 and 1 b-tagged jets are pure enough to study W/Z + light and W/Z + charm
- Z+cc and Z+bb: can be studied in 2-tagged regions in the 2-lepton channel
- W+cc and W+bb: modelling with generator studies

Flavour composition

- Split W/Z+jets according to flavour of two leading jets
- Flavour labelling: cone-based matching ($\Delta R = 0.4$) with hadrons pT>5 GeV

$$V + light,\ V + cl,\ V + \text{heavy flavour} = bb, bc, bl, cc$$

Vcl and Vbb floating freely, normalization determined by global fit.

Process	Scale factor
Wbb	0.83 ± 0.15
Wcl	1.14 ± 0.10
Zbb	1.09 ± 0.05
Zcl	0.88 ± 0.12

ATLAS: W/Z+jets

Cross-section uncertainties

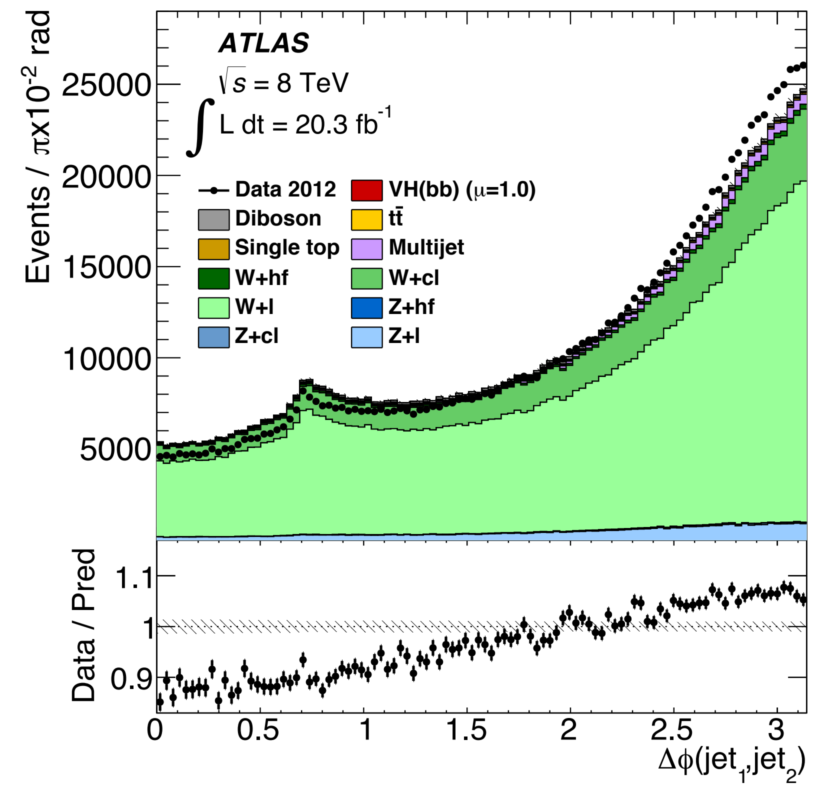
- V+cl and V+bb freely-floating in the fit
- Flavour uncertainties: bl/bb, bc/bb and cc/bb ratio uncertainties from generator comparisons using Alpgen
- 3-to-2 jet ratio:
  - generator comparisons (nominal vs Alpgen)
  - MC-to-data comparison for Z+light

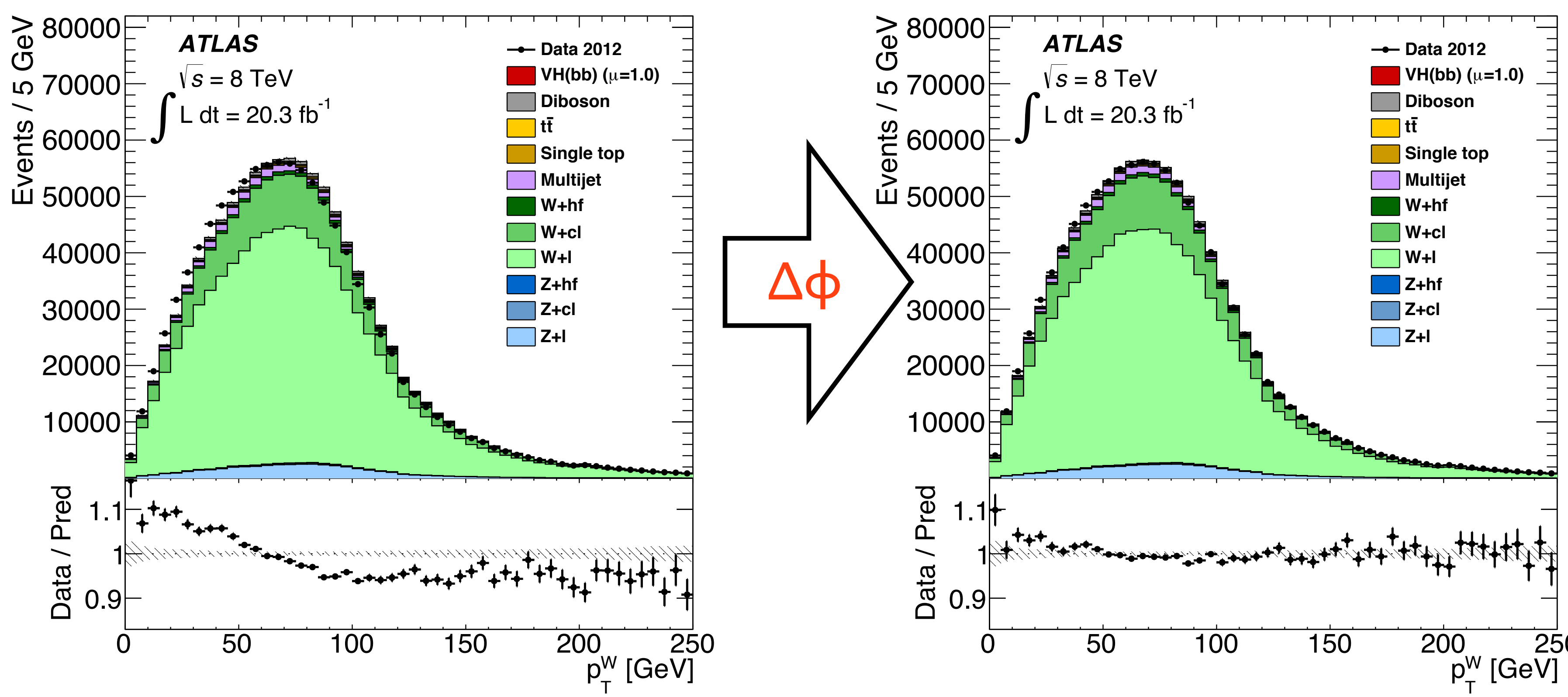
Z+jets	
Zl normalisation, 3/2-jet ratio	5%
Zcl 3/2-jet ratio	26%
Z+hf 3/2-jet ratio	20%
Z+hf/Zbb ratio	12%
$\Delta\phi(\text{jet}_1, \text{jet}_2)$, p_T^V, m_{bb}	S
W+jets	
Wl normalisation, 3/2-jet ratio	10%
Wcl, W+hf 3/2-jet ratio	10%
Wbl/Wbb ratio	35%
Wbc/Wbb, Wcc/Wbb ratio	12%
$\Delta\phi(\text{jet}_1, \text{jet}_2)$, p_T^V, m_{bb}	S

ATLAS: W/Z+jets

Corrections to simulation

- Mismodelling Δϕ between jets in 0 and 1-tag regions
- Correction of W+ll, W+cl and Z+ll processes
- Uncertainty: 50% of the correction when applied, 100% when not applied





ATLAS: W/Z+jets

Z+jets: shape uncertainties and other corrections

- Mismodelling of pT(Z) distribution in 1-tag and 2-tag regions
  - Correction applied to Zcl + Z h.f. processes
  - Uncertainty: 50% of correction on all Z+jets processes
- m_{bb} systematic:
  - derived for Z+jets using data in the 2-tag region (excluding signal region)
  - MC comparisons: Alpgen vs Sherpa difference is covered by data/MC comparison

W+h.f.: shape uncertainties

- Extensive comparisons at generator level:
  - nominal vs Powheg+Pythia8, aMC@NLO+Herwig++ and Alpgen+Herwig
  - scale and PDF variations with aMC@NLO
- Shape uncertainties in pT(W) and m_{bb}

Significant differences: especially m_{bb} shape
$\pm$23% at 50 GeV $\mp$28% at 200 GeV

mainly driven by Sherpa vs Alpgen

ATLAS: impact of bkg uncertainties

After the global fit...

Uncertainties on the background modelling are the ones with the largest impact on the final results:

- W + heavy flavour uncertainties are the most important ones in the analysis
  - dijet mass shape on W+bb and W+cc at high pT(V)
  - Wbl/Wbb normalisation ratio at high pT(V)
  - Wbb background normalisation
  - W+h.f. pT(W) shape on 3-jet category

Lacking good control region.

- Z + heavy flavour shapes and normalisations not far behind, as well as top-pair normalisations

CMS: background modelling

Normalisation uncertainties

- V+jets and top-pair production normalisations determined from data in each channel
- V+b scale factors closer to two:
  - Excess of events with two close-by displaced vertices
  - Interpreted as gluon splitting contribution arXiv:1302.2929, 1310.1349
- Approximately 10% uncertainty on event yields estimated from data

Process	$W(\ell\nu)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$
Low $p_T(V)$			
W + udscg	$1.03 \pm 0.01 \pm 0.05$	–	$0.83 \pm 0.02 \pm 0.04$
W + b	$2.22 \pm 0.25 \pm 0.20$	–	$2.30 \pm 0.21 \pm 0.11$
W + $b\bar{b}$	$1.58 \pm 0.26 \pm 0.24$	–	$0.85 \pm 0.24 \pm 0.14$
Z + udscg	–	$1.11 \pm 0.04 \pm 0.06$	$1.24 \pm 0.03 \pm 0.09$
Z + b	–	$1.59 \pm 0.07 \pm 0.08$	$2.06 \pm 0.06 \pm 0.09$
Z + $b\bar{b}$	–	$0.98 \pm 0.10 \pm 0.08$	$1.25 \pm 0.05 \pm 0.11$
$t\bar{t}$	$1.03 \pm 0.01 \pm 0.04$	$1.10 \pm 0.05 \pm 0.06$	$1.01 \pm 0.02 \pm 0.04$
Intermediate $p_T(V)$			
W + udscg	$1.02 \pm 0.01 \pm 0.07$	–	$0.93 \pm 0.02 \pm 0.04$
W + b	$2.90 \pm 0.26 \pm 0.20$	–	$2.08 \pm 0.20 \pm 0.12$
W + $b\bar{b}$	$1.30 \pm 0.23 \pm 0.14$	–	$0.75 \pm 0.26 \pm 0.11$
Z + udscg	–	–	$1.19 \pm 0.03 \pm 0.07$
Z + b	–	–	$2.30 \pm 0.07 \pm 0.08$
Z + $b\bar{b}$	–	–	$1.11 \pm 0.06 \pm 0.12$
$t\bar{t}$	$1.02 \pm 0.01 \pm 0.15$	–	$0.99 \pm 0.02 \pm 0.03$
High $p_T(V)$			
W + udscg	$1.04 \pm 0.01 \pm 0.07$	–	$0.93 \pm 0.02 \pm 0.03$
W + b	$2.46 \pm 0.33 \pm 0.22$	–	$2.12 \pm 0.22 \pm 0.10$
W + $b\bar{b}$	$0.77 \pm 0.25 \pm 0.08$	–	$0.71 \pm 0.25 \pm 0.15$
Z + udscg	–	$1.11 \pm 0.04 \pm 0.06$	$1.17 \pm 0.02 \pm 0.08$
Z + b	–	$1.59 \pm 0.07 \pm 0.08$	$2.13 \pm 0.05 \pm 0.07$
Z + $b\bar{b}$	–	$0.98 \pm 0.10 \pm 0.08$	$1.12 \pm 0.04 \pm 0.10$
$t\bar{t}$	$1.00 \pm 0.01 \pm 0.11$	$1.10 \pm 0.05 \pm 0.06$	$0.99 \pm 0.02 \pm 0.03$

V+jets split into: udscg, bb, b

Smaller backgrounds taken from simulation:

- Single-top and diboson: 15%
- Consistent with CMS measurements of these processes

CMS: background modelling

Shape uncertainties

- V+jets: uncertainty on shape of BDT distribution
  - Estimated by comparing nominal MC (MadGraph) with Herwig++
- top-pair: uncertainty on shape of BDT distribution
  - Comparing nominal MC (MadGraph) with Powheg and MC@NLO predictions

Combined effect of systematic uncertainties (incl. experimental) is 15% on the expected significance for a SM Higgs

Source	Type	Event yield uncertainty range (%)	Individual contribution to μ uncertainty (%)	Effect of removal on μ uncertainty (%)
Luminosity	norm.	2.2–2.6	<2	<0.1
Lepton efficiency and trigger (per lepton)	norm.	3	<2	<0.1
$Z(\nu\nu)H$ triggers	shape	3	<2	<0.1
Jet energy scale	shape	2–3	5.0	0.5
Jet energy resolution	shape	3–6	5.9	0.7
Missing transverse energy	shape	3	3.2	0.2
b-tagging	shape	3–15	10.2	2.1
Signal cross section (scale and PDF)	norm.	4	3.9	0.3
Signal cross section (p_T boost, EW/QCD)	norm.	2/5	3.9	0.3
Monte Carlo statistics	shape	1–5	13.3	3.6
Backgrounds (data estimate)	norm.	10	15.9	5.2
Single-top-quark (simulation estimate)	norm.	15	5.0	0.5
Dibosons (simulation estimate)	norm.	15	5.0	0.5
MC modeling (V+jets and $t\bar{t}$)	shape	10	7.4	1.1

Summary

Background modelling for VH(bb)

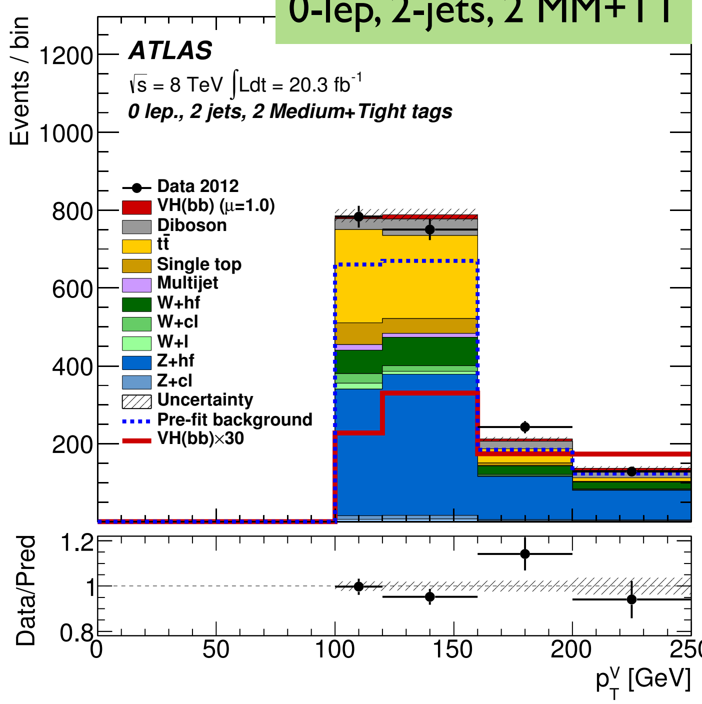
- VH(bb) analysis is extremely challenging:
  - Contaminated by many different and large backgrounds
- Background modelling is a central problem in the VH(bb) searches
  - Normalizations are mainly data-driven
  - Main shape uncertainties come from W+h.f. process: generator comparisons due to lack of good control region

Future (LHC Run 2 and beyond):

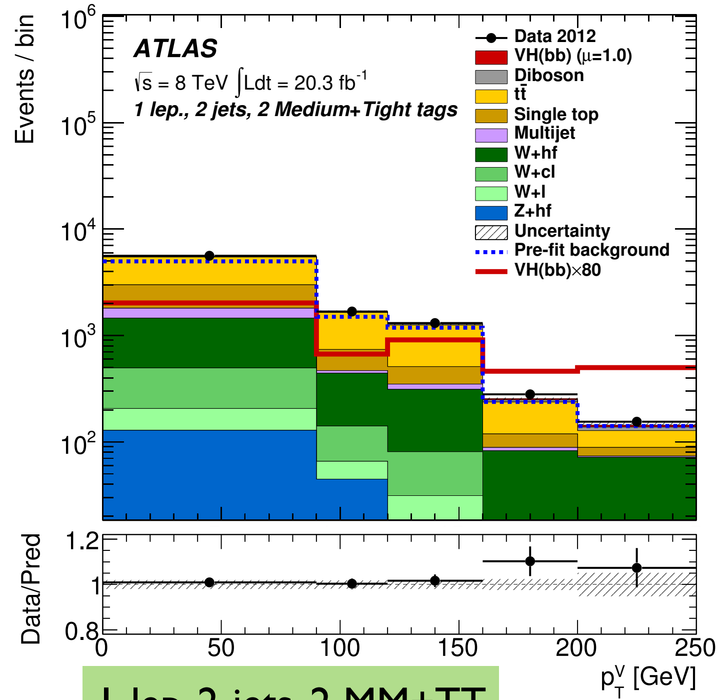
- Essential to get a handle on SM processes (e.g. g→bb)
- Focus on boosted region (highest sensitivity) requires effort on MC modelling

Backup

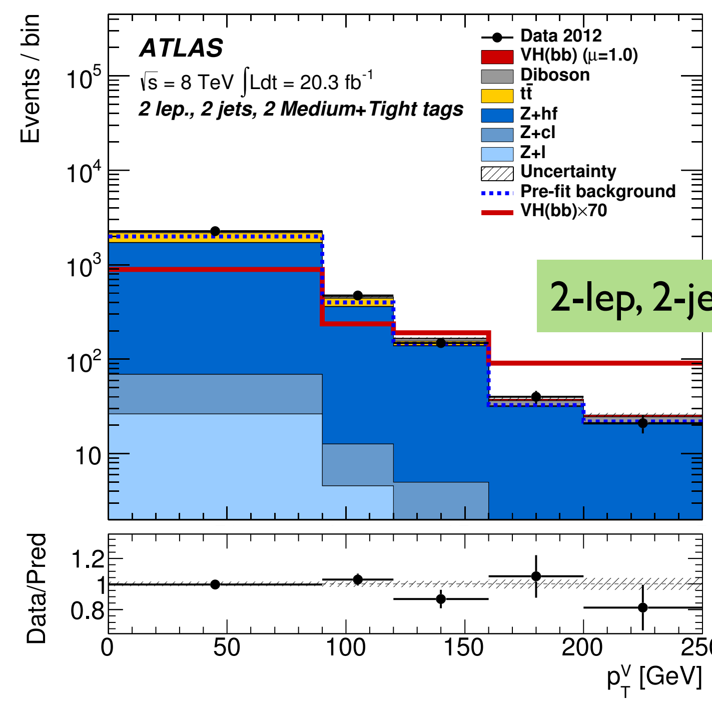
ATLAS: bkg composition example



0-lep, 2-jets, 2 MM+TT



1-lep, 2-jets, 2 MM+TT

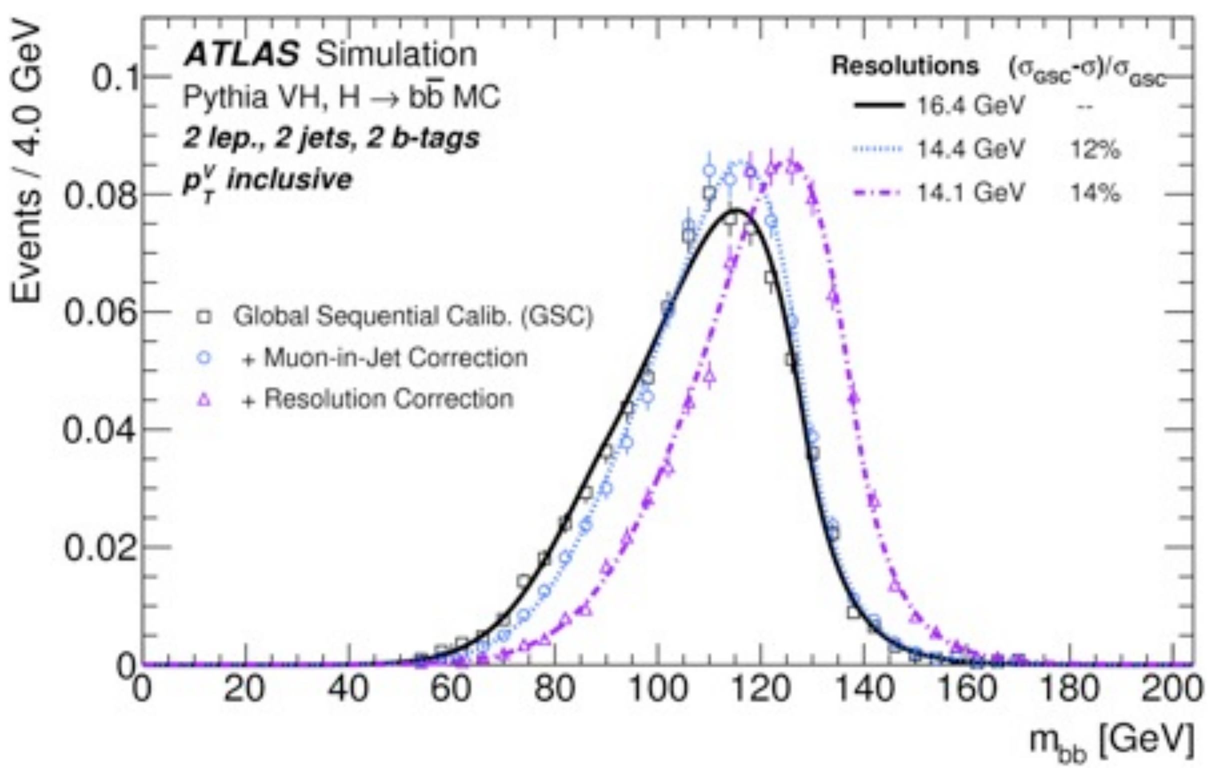


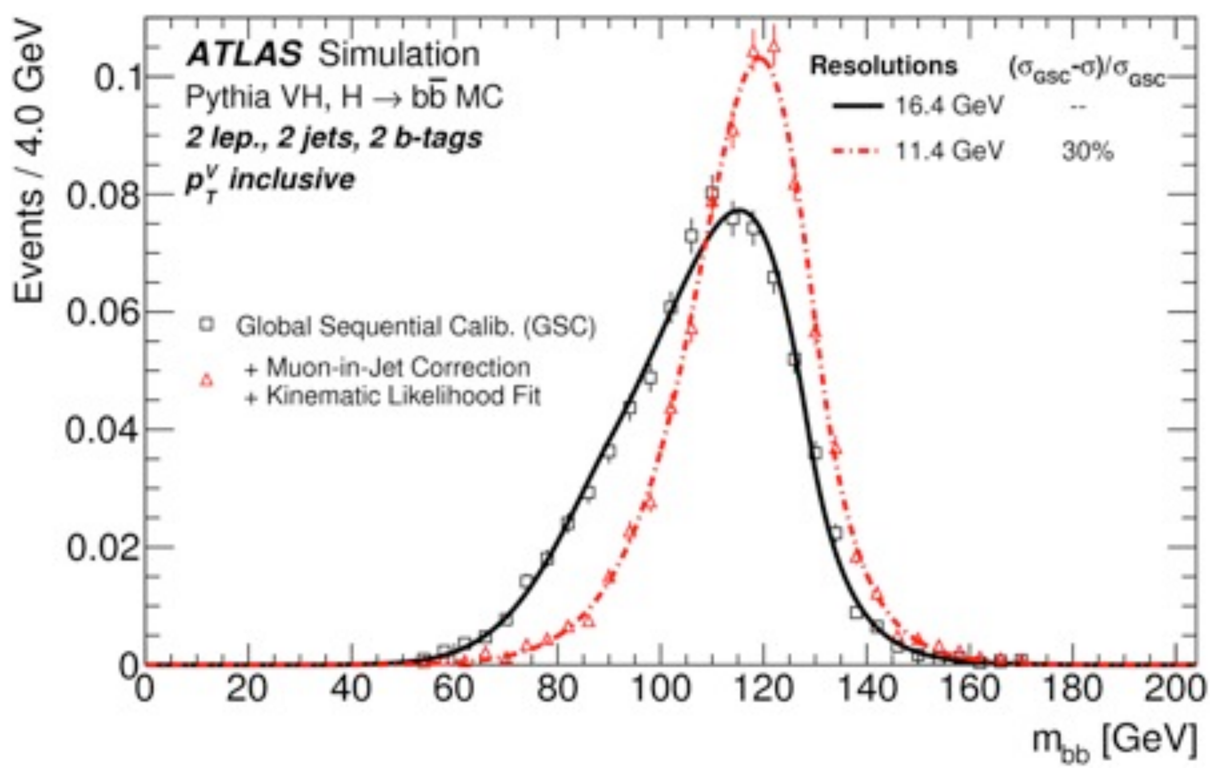
2-lep, 2-jets, 2 MM+TT

ATLAS: mass resolution

- Looking for an excess in m_{bb} distribution (most important variable in MVA)
- Better resolution yields better result (for a given background shape)
- Methods:
  - muon-in-jet correction
  - resolution correction
  - kinematic likelihood fit (2-lepton)

30% improvement on resolution (2-lepton channel)





Event Selection

<table>
<tr><td>Variable</td><td colspan="5">Dijet-mass analysis</td><td colspan="2">Multivariate analysis</td></tr>
<tr><td colspan="8">Common selection</td></tr>
<tr><td>p_{T}^{V} [GeV]</td><td>0–90</td><td>$90^{(*)}$–120</td><td>120–160</td><td>160–200</td><td>> 200</td><td>0–120</td><td>> 120</td></tr>
<tr><td>$\Delta R(\mathrm{jet}_1, \mathrm{jet}_2)$</td><td>0.7–3.4</td><td>0.7–3.0</td><td>0.7–2.3</td><td>0.7–1.8</td><td>< 1.4</td><td colspan="2">> 0.7 (p_{T}^{V}<200 GeV)</td></tr>
<tr><td colspan="8">0-lepton selection</td></tr>
<tr><td>$p_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]</td><td rowspan="5">NU</td><td>> 30</td><td colspan="3">> 30</td><td rowspan="5">NU</td><td>> 30</td></tr>
<tr><td>$\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}})$</td><td>$< \pi/2$</td><td colspan="3">$< \pi/2$</td><td>$< \pi/2$</td></tr>
<tr><td>$\min[\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet})]$</td><td>–</td><td colspan="3">> 1.5</td><td>> 1.5</td></tr>
<tr><td>$\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{dijet})$</td><td>$> 2.2$</td><td colspan="3">$> 2.8$</td><td>–</td></tr>
<tr><td>$\sum_{i=1}^{N_{\mathrm{jet}}=2(3)} p_{\mathrm{T}}^{\mathrm{jet}_i}$ [GeV]</td><td>> 120 (NU)</td><td colspan="3">> 120 (150)</td><td>> 120 (150)</td></tr>
<tr><td colspan="8">1-lepton selection</td></tr>
<tr><td>m_{T}^{W} [GeV]</td><td colspan="5">< 120</td><td colspan="2">–</td></tr>
<tr><td>H_{T} [GeV]</td><td colspan="2">> 180</td><td colspan="3">–</td><td>> 180</td><td>–</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]</td><td colspan="2">–</td><td colspan="2">> 20</td><td>> 50</td><td>–</td><td>> 20</td></tr>
<tr><td colspan="8">2-lepton selection</td></tr>
<tr><td>$m_{\ell\ell}$ [GeV]</td><td colspan="5">83-99</td><td colspan="2">71-121</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]</td><td colspan="5">< 60</td><td colspan="2">–</td></tr>
</table>

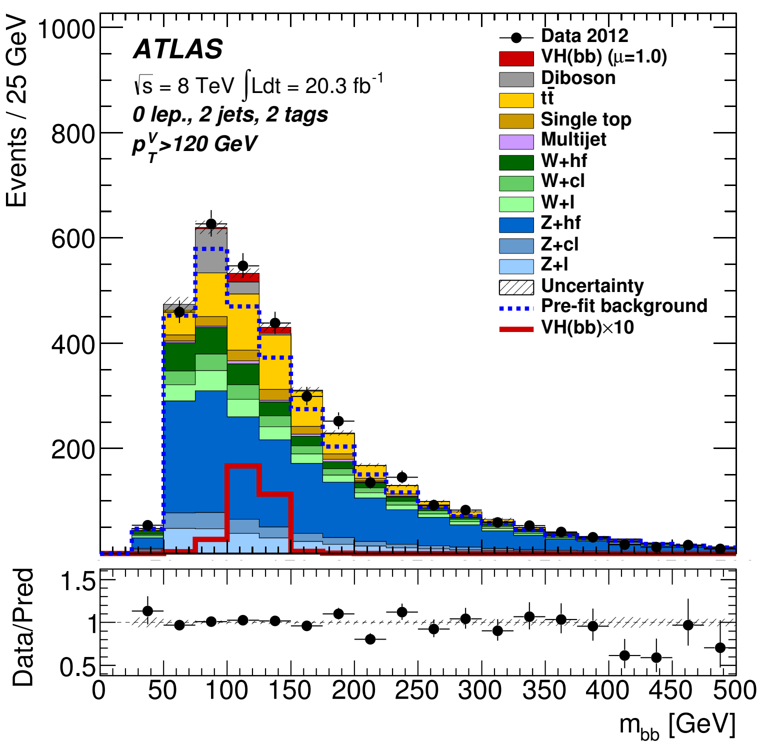
BDT input variables

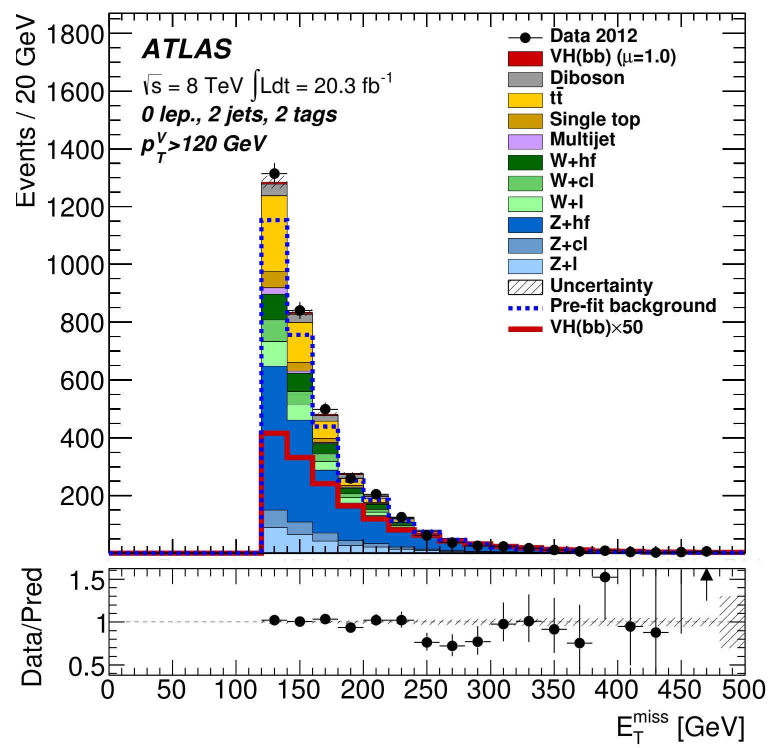
Variable	0-Lepton	1-Lepton	2-Lepton
p_T^V		×	×
E_T^{miss}	×	×	×
$p_T^{b_1}$	×	×	×
$p_T^{b_2}$	×	×	×
m_{bb}	×	×	×
$\Delta R(b_1, b_2)$	×	×	×
$\|\Delta\eta(b_1, b_2)\|$	×		×
$\Delta\phi(V, bb)$	×	×	×
$\|\Delta\eta(V, bb)\|$			×
H_T	×		
$\min[\Delta\phi(\ell, b)]$		×	
m_T^W		×	
$m_{\ell\ell}$			×
$MV1c(b_1)$	×	×	×
$MV1c(b_2)$	×	×	×
	Only in 3-jet events		
$p_T^{jet_3}$	×	×	×
m_{bbj}	×	×	×

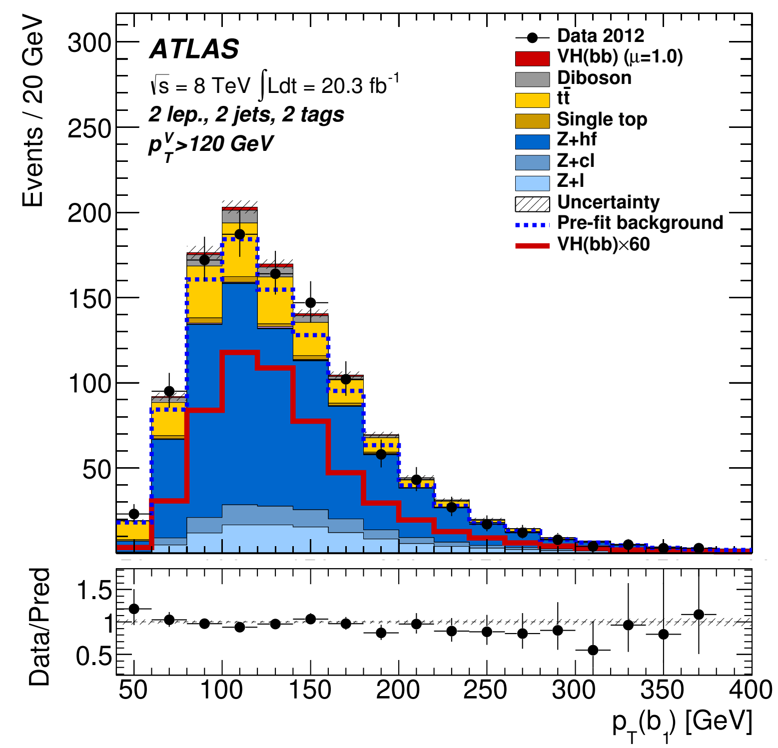
Systematic uncertainties

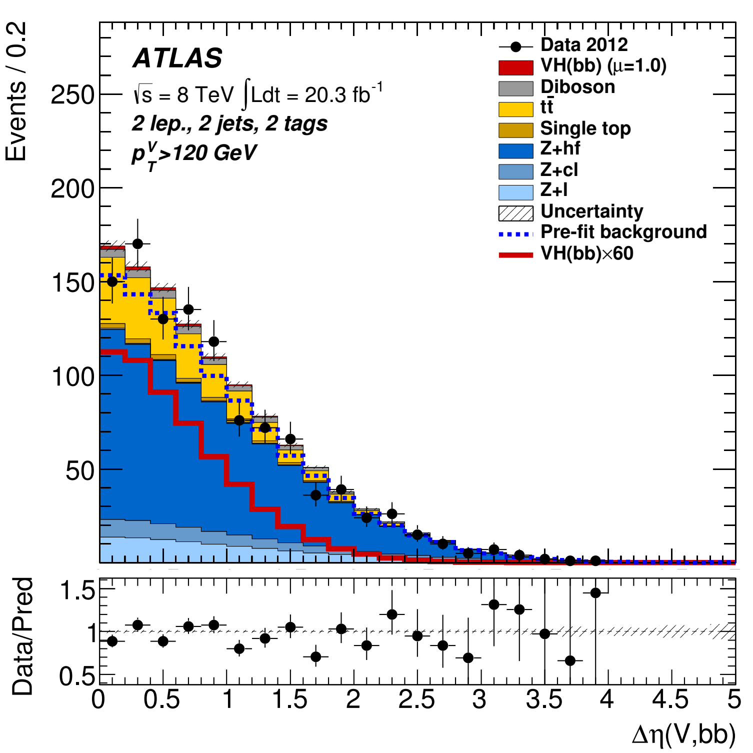
Signal	
Cross section (scale)	1% ($q\bar{q}$), 50% (gg)
Cross section (PDF)	2.4% ($q\bar{q}$), 17% (gg)
Branching ratio	3.3 %
Acceptance (scale)	1.5%–3.3%
3-jet acceptance (scale)	3.3%–4.2%
p_{T}^{V} shape (scale)	S
Acceptance (PDF)	2%–5%
p_{T}^{V} shape (NLO EW correction)	S
Acceptance (parton shower)	8%–13%
Z+jets	
Zl normalisation, 3/2-jet ratio	5%
Zcl 3/2-jet ratio	26%
Z+hf 3/2-jet ratio	20%
Z+hf/Zbb ratio	12%
$\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$, p_{T}^{V}, m_{bb}	S
W+jets	
Wl normalisation, 3/2-jet ratio	10%
Wcl, W+hf 3/2-jet ratio	10%
Wbl/Wbb ratio	35%
Wbc/Wbb, Wcc/Wbb ratio	12%
$\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$, p_{T}^{V}, m_{bb}	S
$t\bar{t}$	
3/2-jet ratio	20%
High/low-p_{T}^{V} ratio	7.5%
Top-quark p_{T}, m_{bb}, $E_{\mathrm{T}}^{\mathrm{miss}}$	S
Single top	
Cross section	4% (s-,t-channel), 7% (Wt)
Acceptance (generator)	3%–52%
m_{bb}, $p_{\mathrm{T}}^{b_1}$	S
Diboson	
Cross section and acceptance (scale)	3%–29%
Cross section and acceptance (PDF)	2%–4%
m_{bb}	S
Multijet	
0-, 2-lepton channels normalisation	100%
1-lepton channel normalisation	2%–60%
Template variations, reweighting	S

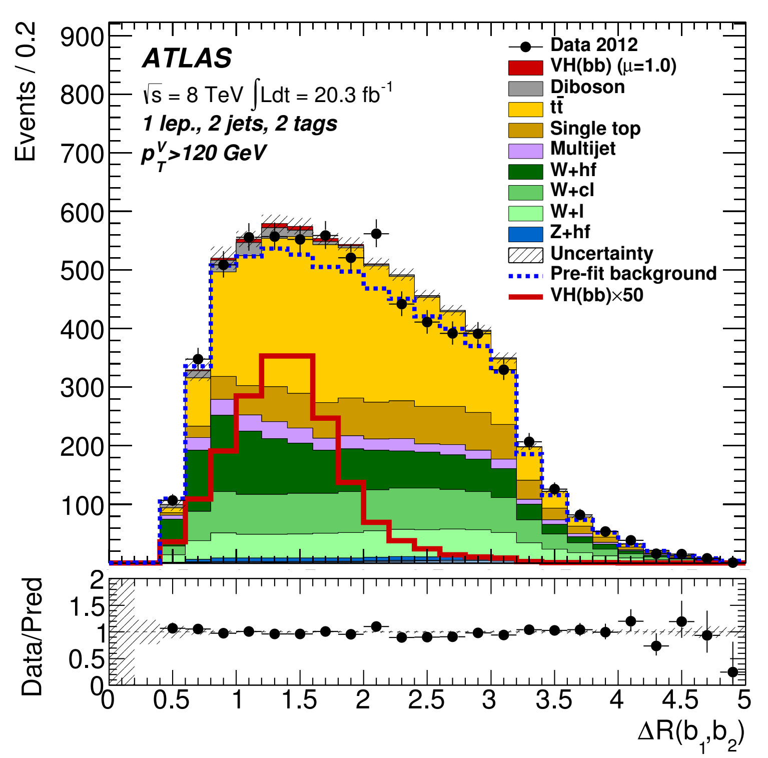
BDT input variables

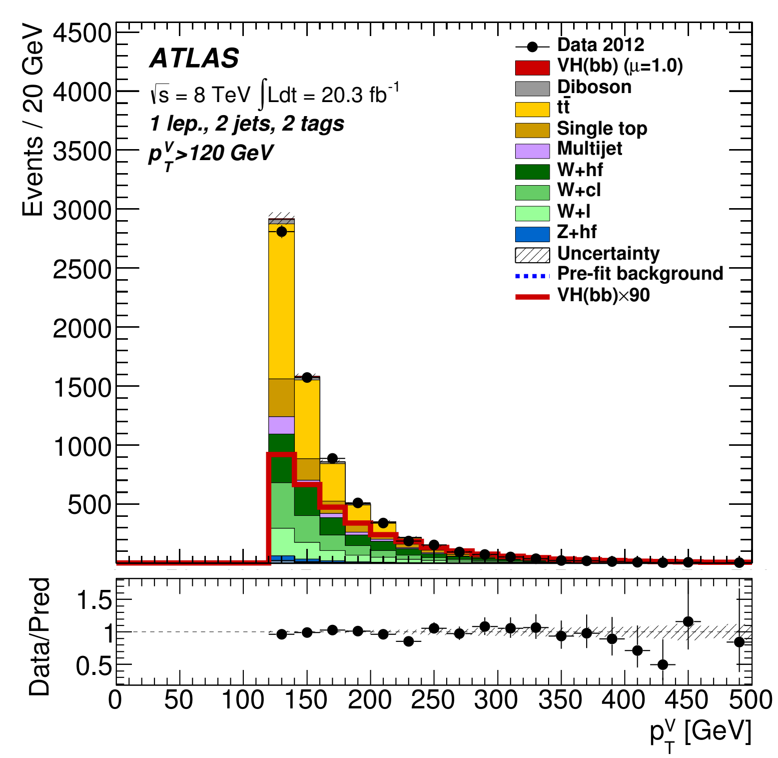




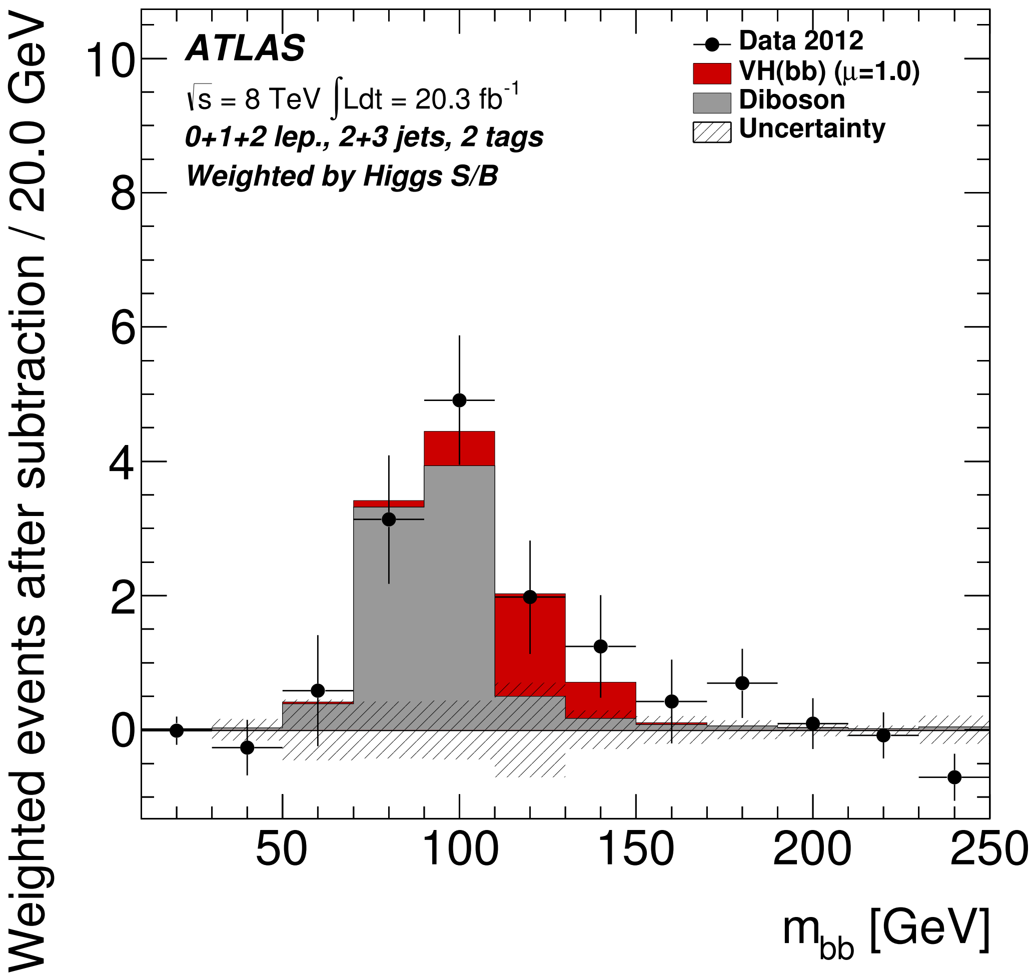


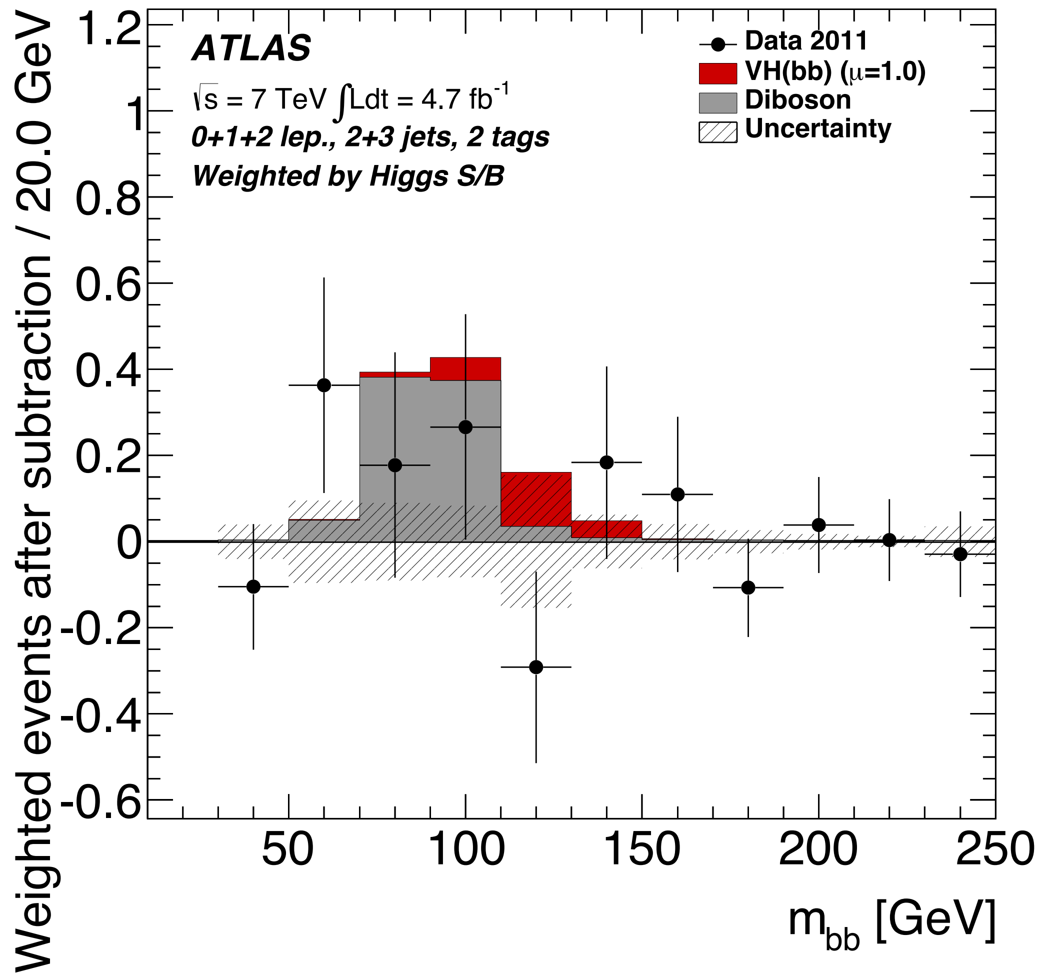






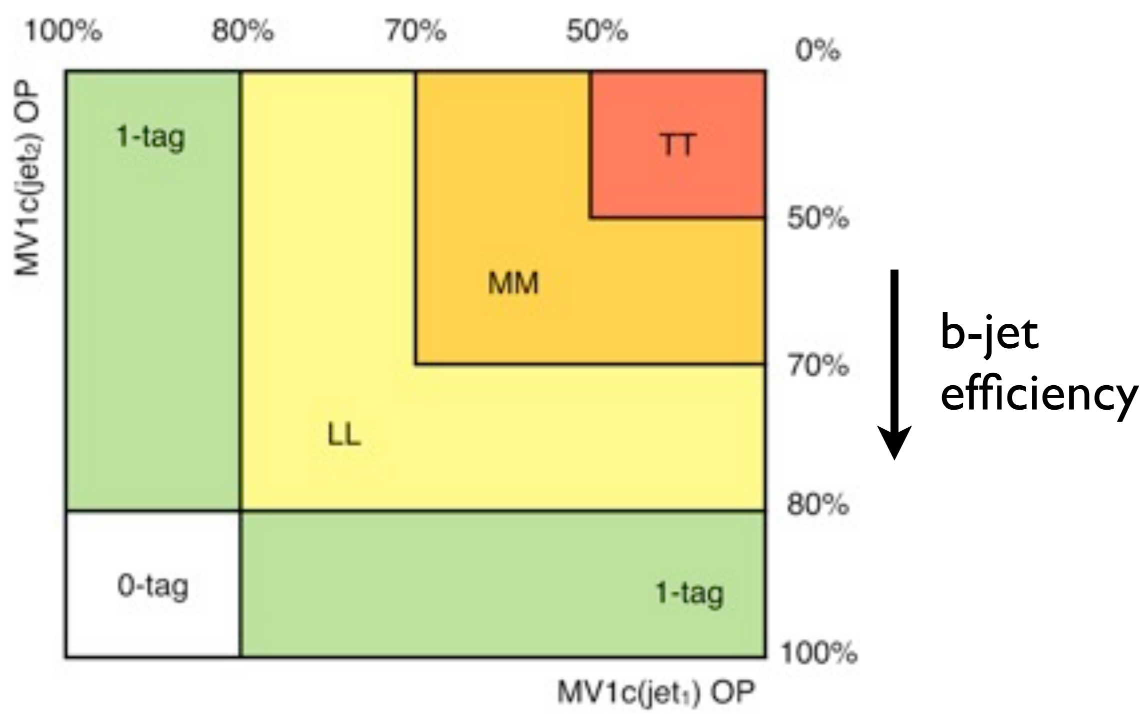
ATLAS: VZ(bb) @ 7 and 8 TeV

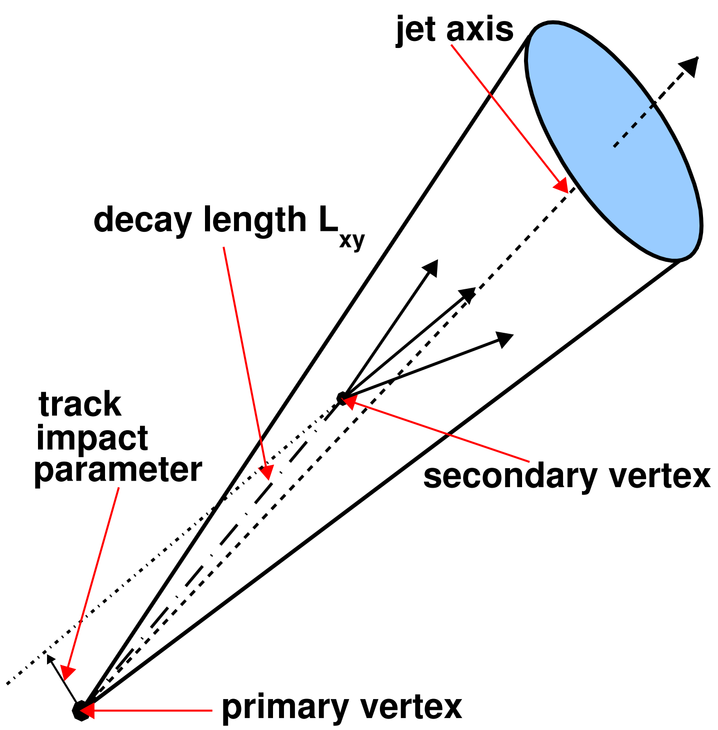




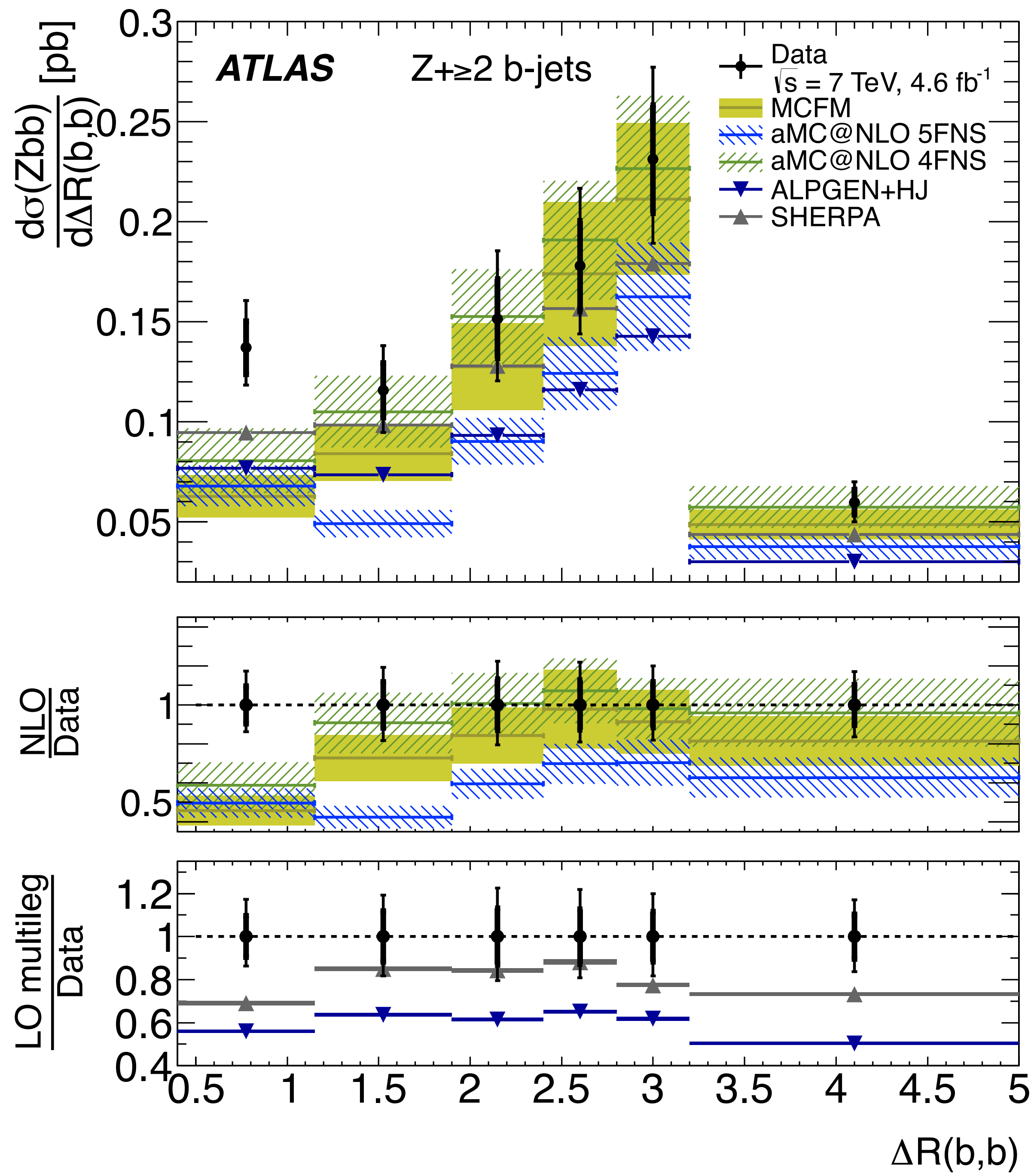
b-tagging

- Reconstruct Higgs boson candidate with two b-tagged jets
- b-tagging algorithm used to identify jets originating from b-quarks
  - Combining information on track impact parameters, secondary vertices and b/c decay chains in a neural network
- Use different b-tagging operating points to define event categories
  - Allows for different charm rejection rates: handle on V+jets processes





Z+bb @ 7 TeV


ATLAS
Z+≥2 b-jets
Data
√s = 7 TeV, 4.6 fb⁻¹
MCFM
aMC@NLO 5FNS
aMC@NLO 4FNS
ALPGEN+HJ
SHERPA
dσ(Zbb)/dΔR(b,b) [pb]
NLO/Data
LO multileg/Data
ΔR(b,b)

Event Selection - signal region

Variable	W($\ell\nu$)H	W($\tau\nu$)H	Z($\ell\ell$)H	Z($\nu\nu$)H
p_T(V)	[100–130] [130–180] [>180]	[>120]	[50–100] [>100]	[100–130] [130–170] [>170]
$m_{\ell\ell}$	–	–	[75–105]	–
$p_T(j_1)$	>30	>30	>20	>60
$p_T(j_2)$	>30	>30	>20	>30
p_T(jj)	>100	>120	–	[>100] [>130] [>130]
m(jj)	<250	<250	[40–250] [< 250]	<250
E_T^{miss}	>45	>80	–	[100–130] [130–170] [> 170]
$p_T(\tau)$	–	>40	–	–
p_T(track)	–	>20	–	–
CSV_{max}	>0.40	>0.40	[>0.50] [>0.244]	>0.679
CSV_{min}	>0.40	>0.40	>0.244	>0.244
N_{aj}	–	–	–	[< 2] [–] [–]
$N_{a\ell}$	=0	=0	–	=0
$\Delta\phi$(V,H)	–	–	–	>2.0
$\Delta\phi(E_T^{miss}$, jet)	–	–	–	[>0.7] [>0.7] [>0.5]
$\Delta\phi(E_T^{miss}, E_T^{miss}$(tracks))	–	–	–	<0.5
E_T^{miss} significance	–	–	–	[>3] [–] [–]
$\Delta\phi(E_T^{miss}, \ell)$	$< \pi/2$	–	–	–

Event Selection - control regions

W(μν)H, W(eν)H channels

Variable	W+LF	$t\bar{t}$	W+HF
$p_T(j_1)$	>30	>30	>30
$p_T(j_2)$	>30	>30	>30
$p_T(jj)$	>100	>100	>100
$m(jj)$	<250	<250	<250, $\notin$[90-150]
CSV_{max}	$\in$[0.244–0.898]	>0.898	>0.898
N_{aj}	<2	>1	=0
$N_{a\ell}$	=0	=0	=0
E_T^{miss}	>45	>45	>45
E_T^{miss} significance	>2.0(μ) >3.0(e)	–	–

Z(ll)H channel

Variable	Z+jets	$t\bar{t}$
$m_{\ell\ell}$	[75–105]	$\notin$[75–105]
$p_T(j_1)$	>20	>20
$p_T(j_2)$	>20	>20
$p_T(V)$	>50	[50–100]
$m(jj)$	<250, $\notin$[80–150]	<250, $\notin$[80–150]
CSV_{max}	>0.244	>0.244
CSV_{min}	>0.244	>0.244

Z(νν)H channel

Variable	Z+LF	Z+HF	$t\bar{t}$	W+LF	W+HF
E_T^{miss}	[100–130] [130–170] [>170]	[100–130] [130–170] [>170]	[100–130] [130–170] [>170]	[100–130] [130–170] [>170]	[100–130] [130–170] [> 170]
$p_T(j_1)$	>60	>60	>60	>60	>60
$p_T(j_2)$	>30	>30	>30	>30	>30
$p_T(jj)$	[> 100] [> 130] [> 130]	[> 100] [> 130] [> 130]	[> 100] [> 130] [> 130]	[> 100] [> 130] [> 130]	[> 100] [> 130] [> 130]
$m(jj)$	<250	<250, $\notin$[100–140]	<250, $\notin$[100–140]	<250	<250, $\notin$[100–140]
CSV_{max}	[0.244 – 0.898]	>0.679	>0.898	[0.244 – 0.898]	>0.679
CSV_{min}	–	>0.244	–	–	>0.244
N_{aj}	[<2] [–] [–]	[<2] [–] [–]	≥ 1	=0	=0
$N_{a\ell}$	=0	=0	=1	=1	=1
$\Delta\phi(V,H)$	–	>2.0	–	–	>2.0
$\Delta\phi(E_T^{miss}, jet)$	[>0.7] [>0.7] [>0.5]	[>0.7] [>0.7] [>0.5]	[>0.7] [>0.7] [>0.5]	[>0.7] [>0.7] [>0.5]	[>0.7] [>0.7] [>0.5]
$\Delta\phi(E_T^{miss}, E_T^{miss}(tracks))$	<0.5	<0.5	–	–	–
E_T^{miss} significance	[>3] [–] [–]	[>3] [–] [–]	[>3] [–] [–]	[>3] [–] [–]	[>3] [–] [–]

BDT input variables

showing all four subsets