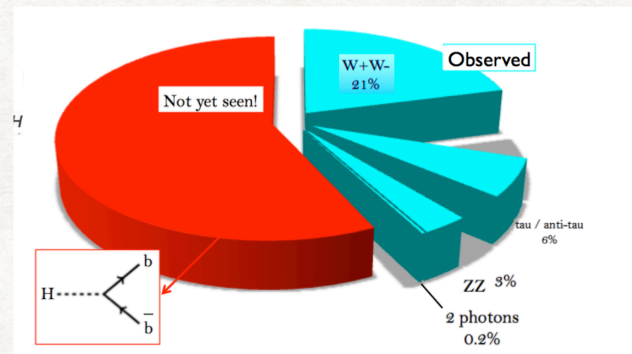
Measurement of VH(bb) with boosted topology at ATLAS

Cheng Chen
on behalf of the ATLAS collaboration

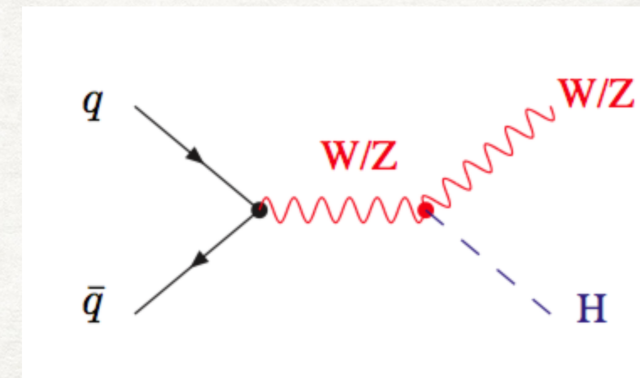
on CLHCP
07 November 2020

Review of VH(bb) analysis

- ## Motivation of VH(bb)
  - Direct evidence for Higgs Yukawa coupling
  - Largest branching ratio of Higgs decays
  - Requiring VH production to suppress multi-jets background

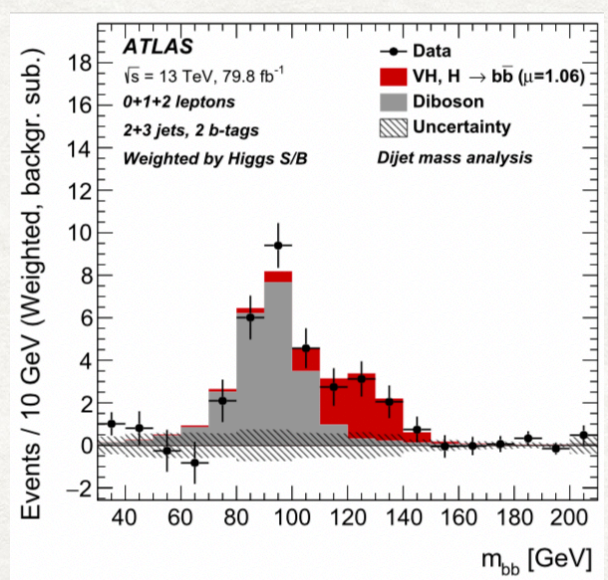


- Branching ratio H→bb

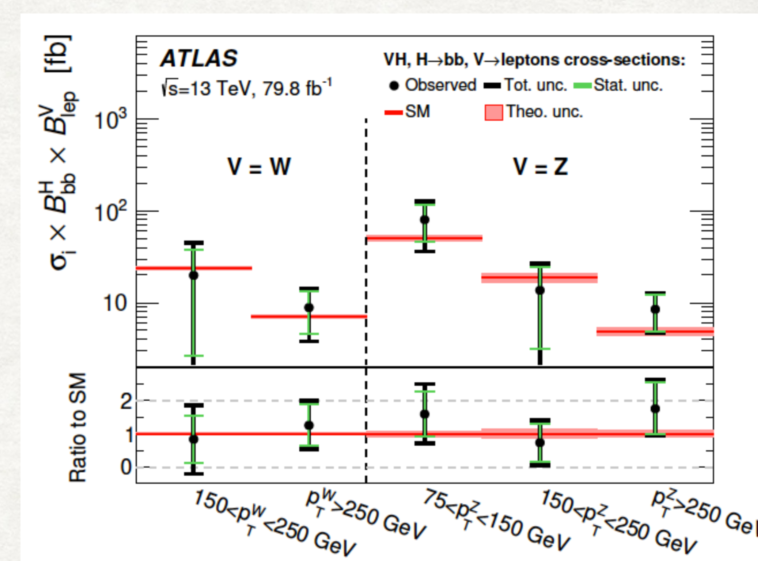


- VH production

- ## Previous VH(bb) results in ATLAS
  - Observation of H(bb) and VH [Phys. Lett. B 786 (2018) 59]
  - STXS measurement [JHEP 05 (2019) 141]



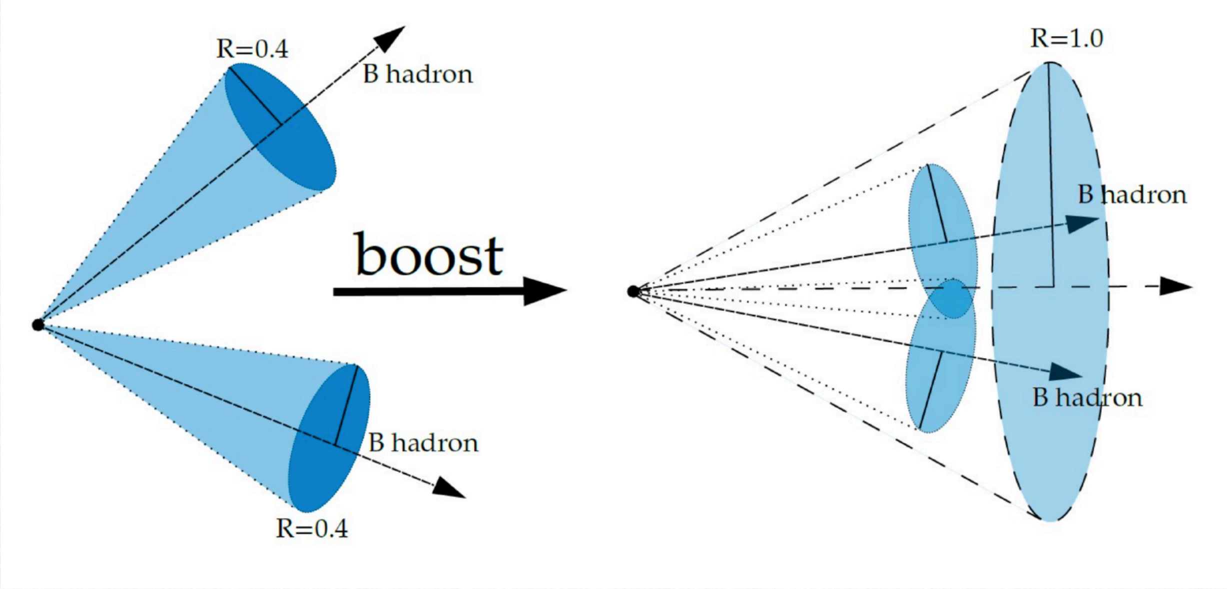
- Observation of Hbb and VH



- STXS measurement

Introduction of VH(bb) boosted analysis

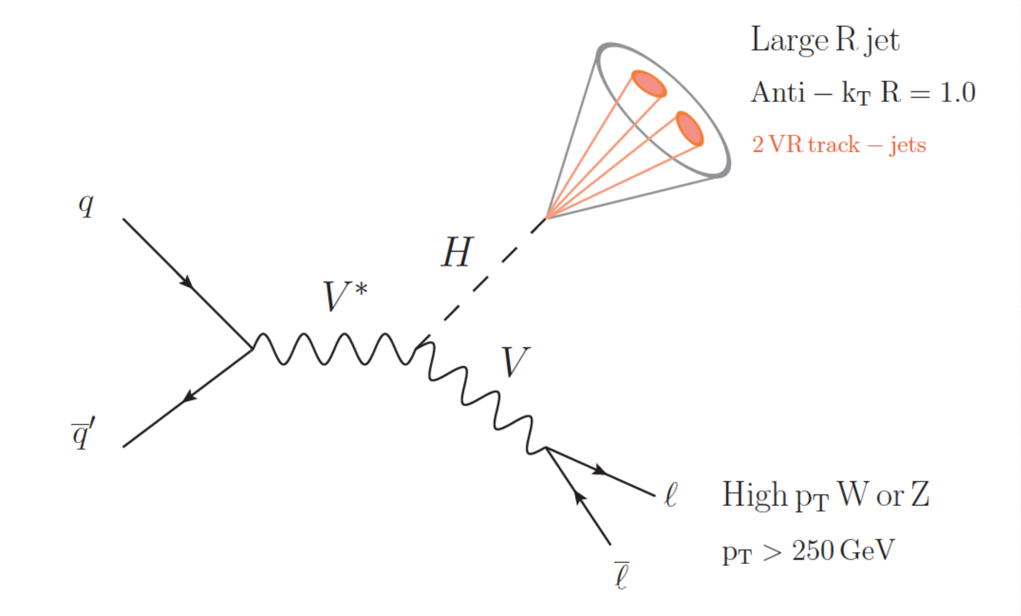
- Motivation
  - Exploit the higher Higgs boson p_T regime
  - Search for deviations from the SM



- First VH(bb) boosted analysis [submitted to PLB, arXiv:2008.02508]
  - Using 139 fb^{-1} of pp collision data collected by ATLAS in 2015 to 2018
  - Exploit higher p_T regimes with a simple and robust analysis
  - Understanding usage of boosted reconstruction techniques for VH(bb)

Analysis strategy

- Higgs candidate
  - Large-R jet (R=1.0) with pT > 250 GeV
  - ≥ 2 small-R track jets matched to Large-R jet
  - Leading 2 track jets b-tagged (backup)
  - Mass of Large-R jet (m_J) is considered as the final discriminant



- Three decay channels w.r.t. the vector boson decay modes
  - 0 lepton: Z→νν   1 lepton: W→lν   2 lepton: Z→ll
- Combined likelihood fit is performed to estimate the significance and signal strength (μ)

$$\mu = \frac{(\sigma * BR)_{Obs}}{(\sigma * BR)_{SM}}$$

- Additional validation of the robustness: VZ analysis (backup)
- STXS measurement and EFT interpretation of the results

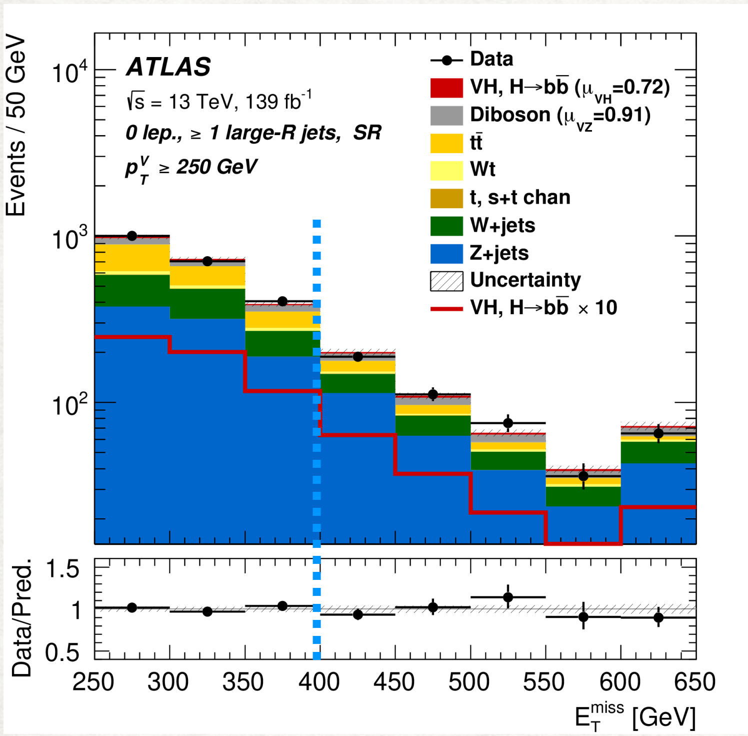
Event selection

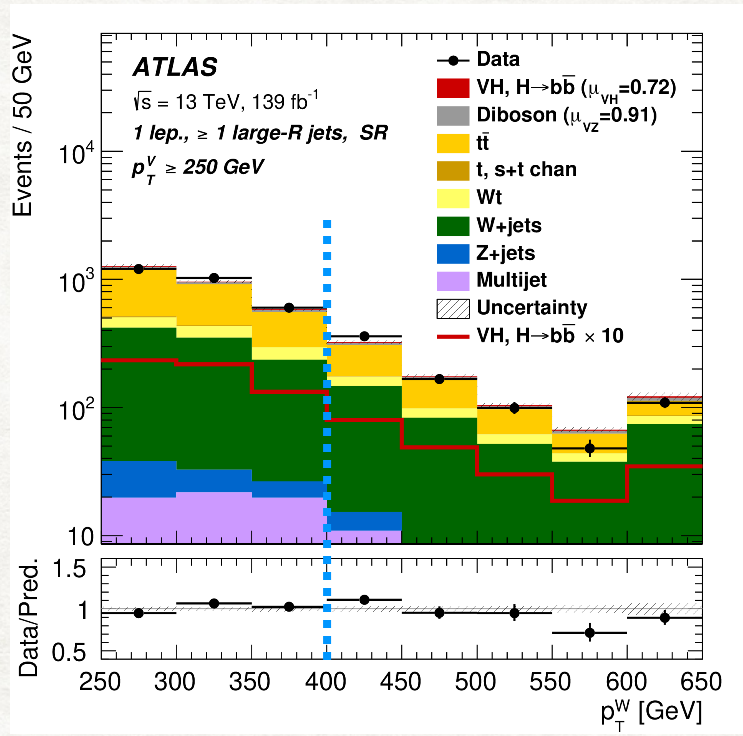
Selection	0 lepton channel	1 lepton channel		2 leptons channel	
		e sub-channel	μ sub-channel	e sub-channel	μ sub-channel
Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single electron	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single electron	$E_{\mathrm{T}}^{\mathrm{miss}}$
Leptons	0 *baseline* leptons	1 *signal* lepton $p_{\mathrm{T}} > 27$ GeV no second *baseline* lepton	$p_{\mathrm{T}} > 25$ GeV	2 *baseline* leptons among which ≥ 1 *signal* lepton, $p_{\mathrm{T}} > 27$ GeV both leptons of the same flavour -	opposite sign muons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 250 GeV	> 50 GeV	-	-	
p_{T}^{V}	$p_{\mathrm{T}}^{V} > 250$ GeV				
Large-R jets	at least one large-R jet, $p_{\mathrm{T}} > 250$ GeV, $\lvert\eta\rvert < 2.0$				
Track-jets	at least two track-jets, $p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.5$, matched to the leading large-R jet				
b-jets	leading two track-jets matched to the leading large-R must be b-tagged (MV2c10, 70%)				
m_{J}	> 50 GeV				
$\min[\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \text{small-}R \text{ jets})]$	$> 30°$	-			
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, H_{\mathrm{cand}})$	$> 120°$	-			
$\Delta\phi\,(E_{\mathrm{T}}^{\mathrm{miss}}, E_{\mathrm{T,\,trk}}^{\mathrm{miss}})$	$< 90°$	-			
$\Delta y(V, H_{\mathrm{cand}})$	-	$\lvert\Delta y(V, H_{\mathrm{cand}})\rvert < 1.4$			
$m_{\ell\ell}$	-			$66\ \mathrm{GeV} < m_{\ell\ell} < 116$ GeV	
Lepton p_{T} imbalance	-			$(p_{\mathrm{T}}^{\ell_1} - p_{\mathrm{T}}^{\ell_2})/p_{\mathrm{T}}^{Z} < 0.8$	

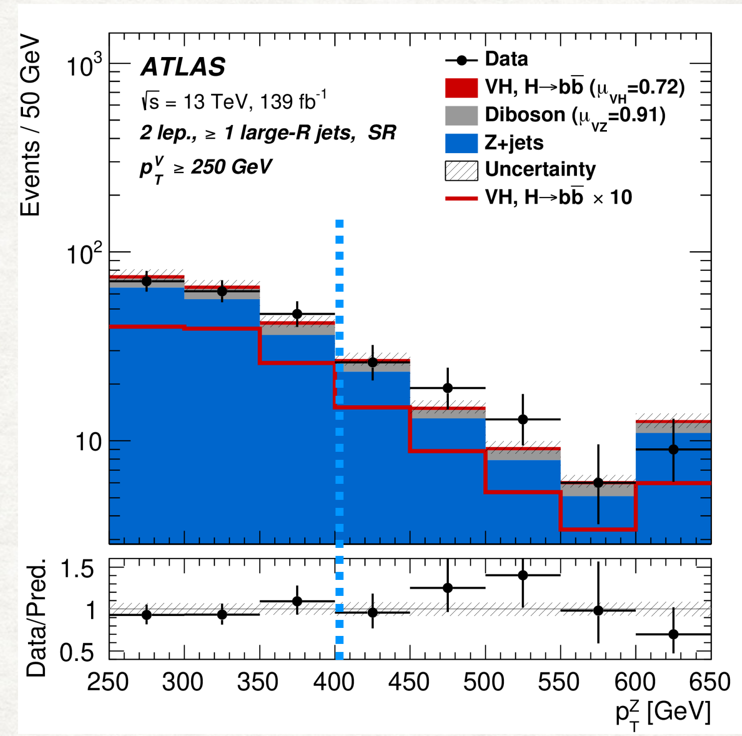
- Vector boson selections (mainly inherited from resolved analysis)
- Higgs boson candidate selections
- Channel specific selections to enhance sensitivity

Event Categorisation

Vector boson p_T





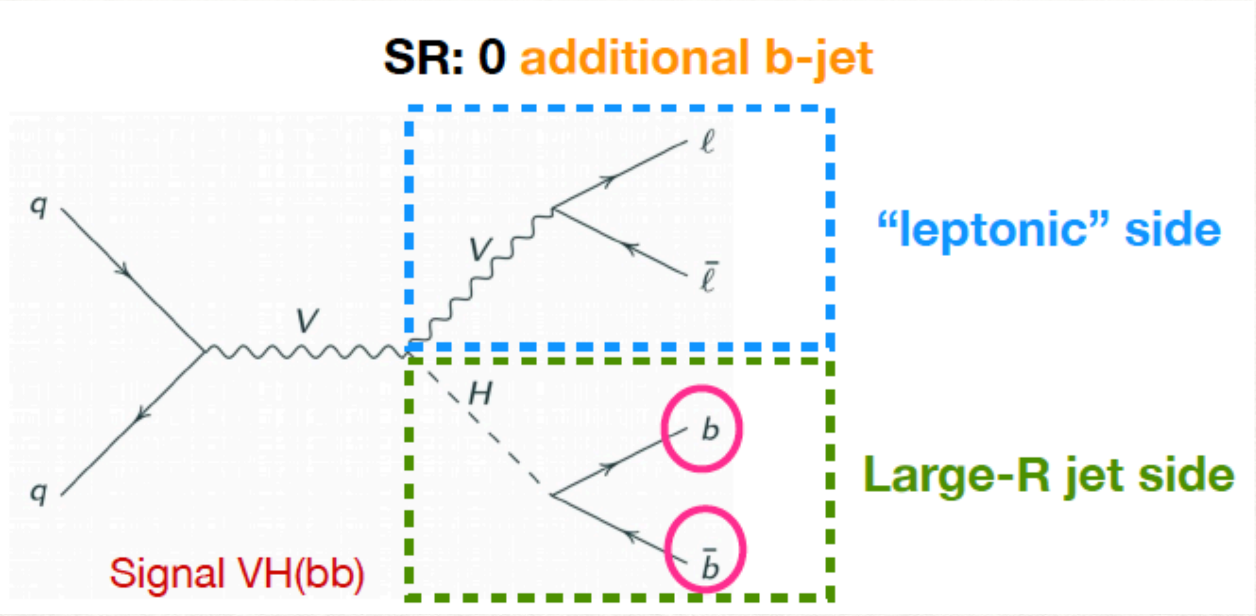


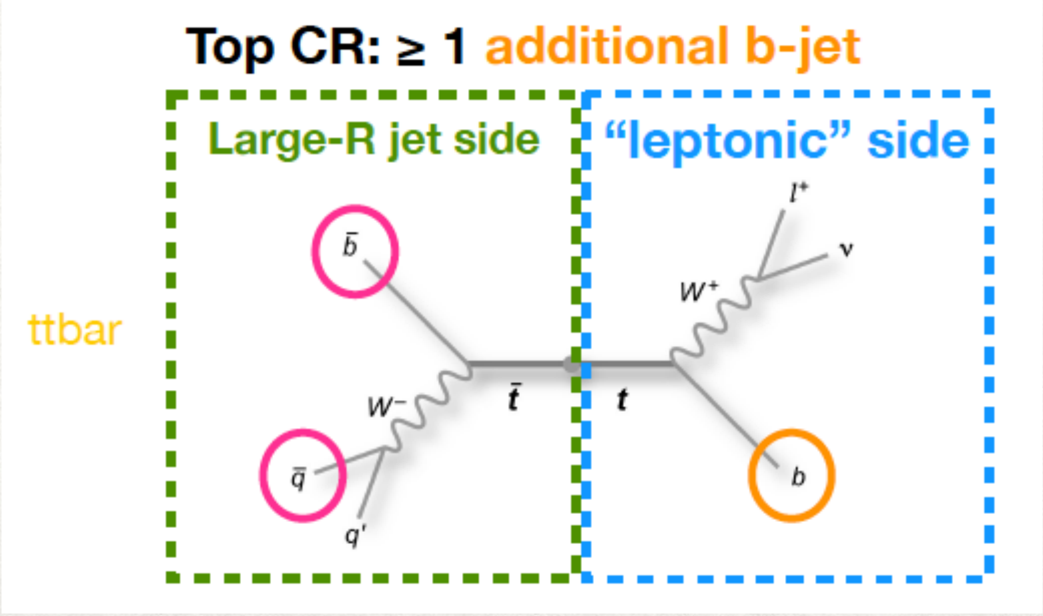
- 2 categories according to vector boson p_T
  - 250-400 GeV
  - ≥ 400 GeV
- This categorisation is tailored towards a differential cross section measurement

Event Categorisation

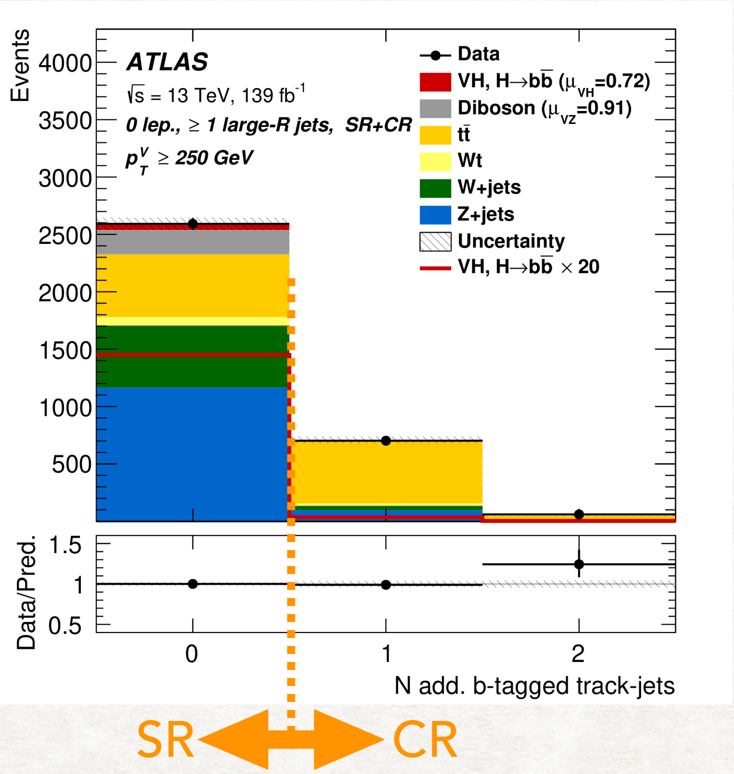
Signal Region v.s Top Control Region in 0/1L

- SR: 0 additional b-jet
- Top CR: ≥1 additional b-jet

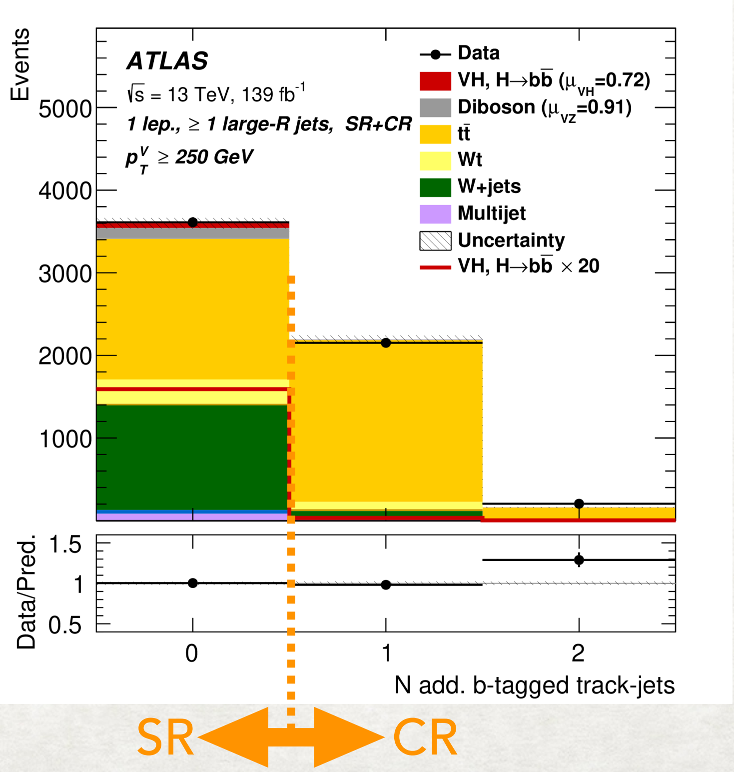




0L



1L

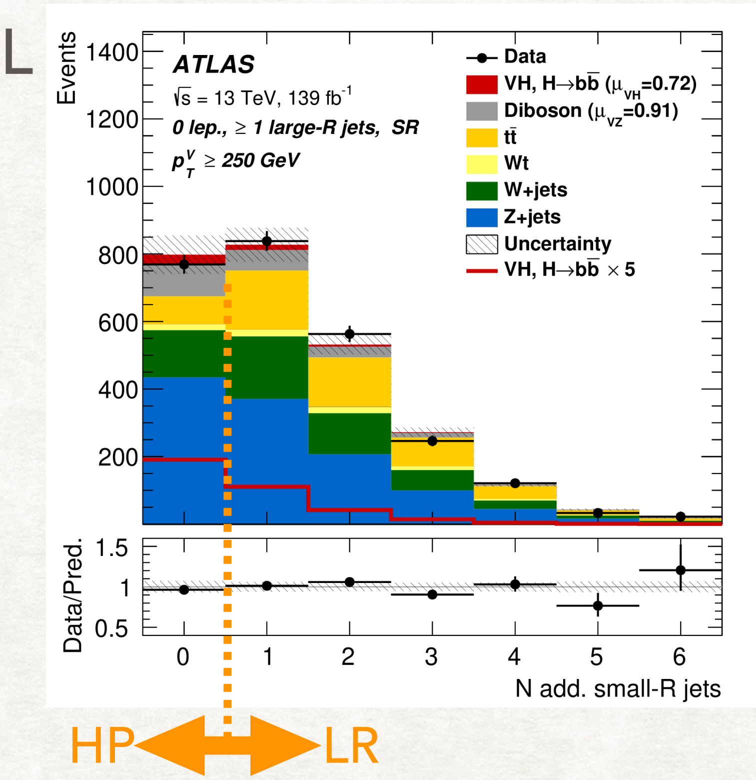


Event Categorisation

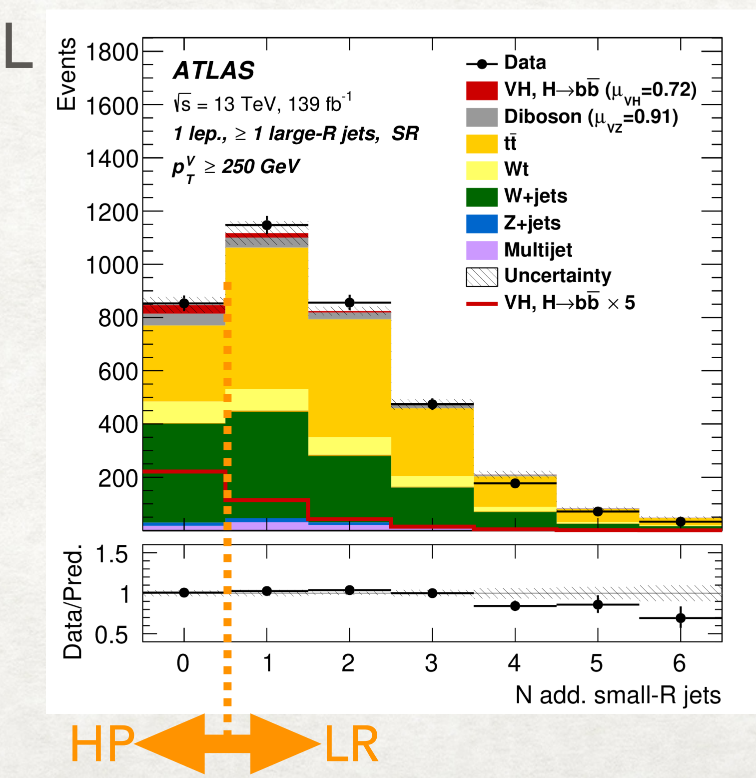
High Purity SR (HP) vs Low Purity SR (LP)

- 0L and 1L only
- Additional small-R jet:
  calo-jets not matched to Large-R jet
- Categorisation
  - HP: 0 additional small-R jets
  - LP: ≥1 additional small-R jets

0L



1L



Event Categorisation

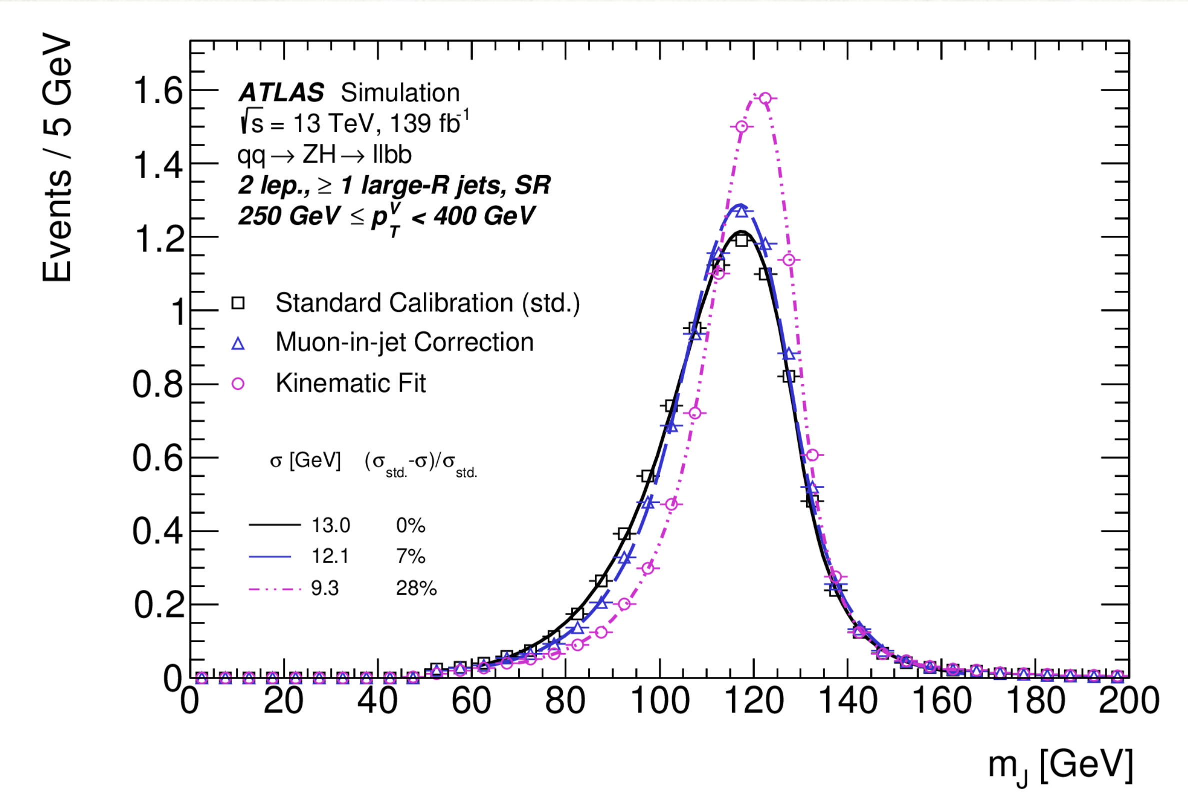
Summary

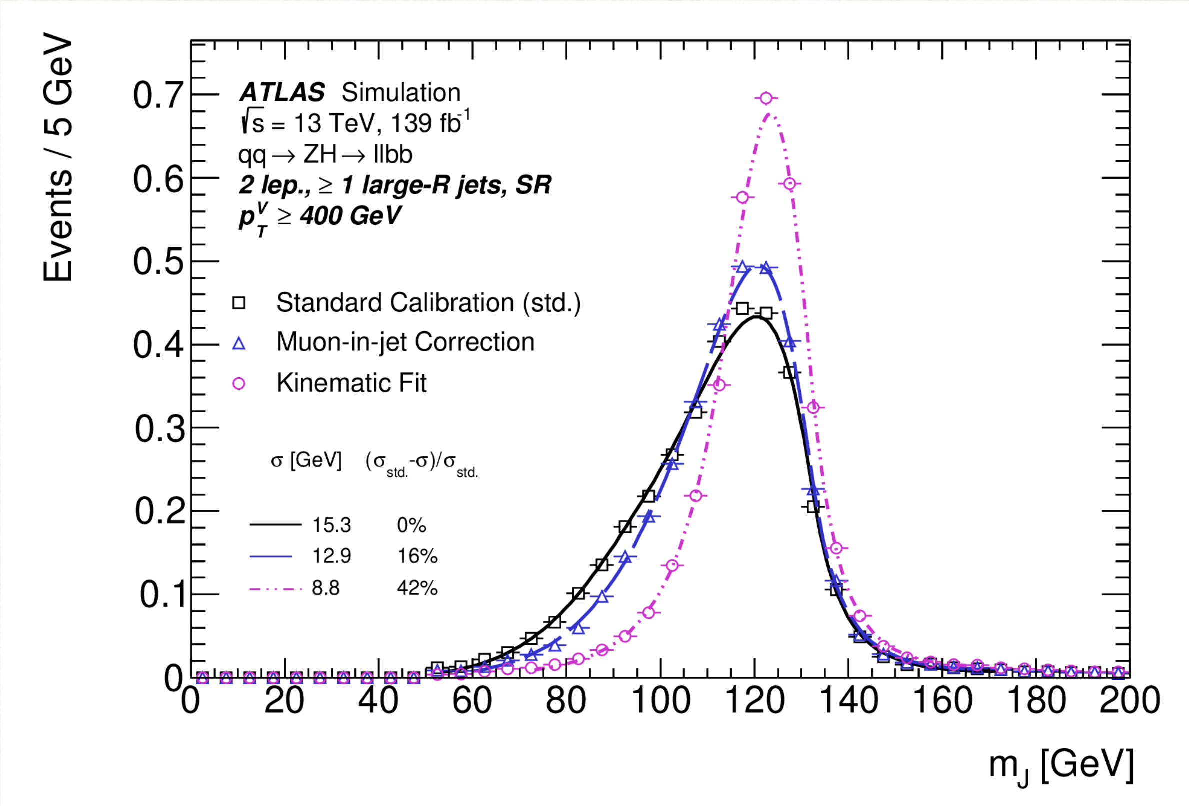
<table>
<tr><th rowspan="4">Channel</th><th colspan="6">Categories</th></tr>
<tr><th colspan="3">$250 < p_T^V < 400$ GeV</th><th colspan="3">$p_T^V \geq 400$ GeV</th></tr>
<tr><th colspan="2">0 add. b-track-jets</th><th rowspan="2">≥ 1 add. b-track-jets</th><th colspan="2">0 add. b-track-jets</th><th rowspan="2">≥ 1 add. b-track-jets</th></tr>
<tr><th>0 add. small-R jets</th><th>≥ 1 add. small-R jets</th><th>0 add. small-R jets</th><th>≥ 1 add. small-R jets</th></tr>
<tr><td>0-lepton</td><td>HP SR</td><td>LP SR</td><td>CR</td><td>HP SR</td><td>LP SR</td><td>CR</td></tr>
<tr><td>1-lepton</td><td>HP SR</td><td>LP SR</td><td>CR</td><td>HP SR</td><td>LP SR</td><td>CR</td></tr>
<tr><td>2-lepton</td><td colspan="3">SR</td><td colspan="3">SR</td></tr>
</table>

- 14 Regions taken into account in the statistical analysis
  - 10 SRs
  - 4 CRs

Large-R jet mass

The final discriminant

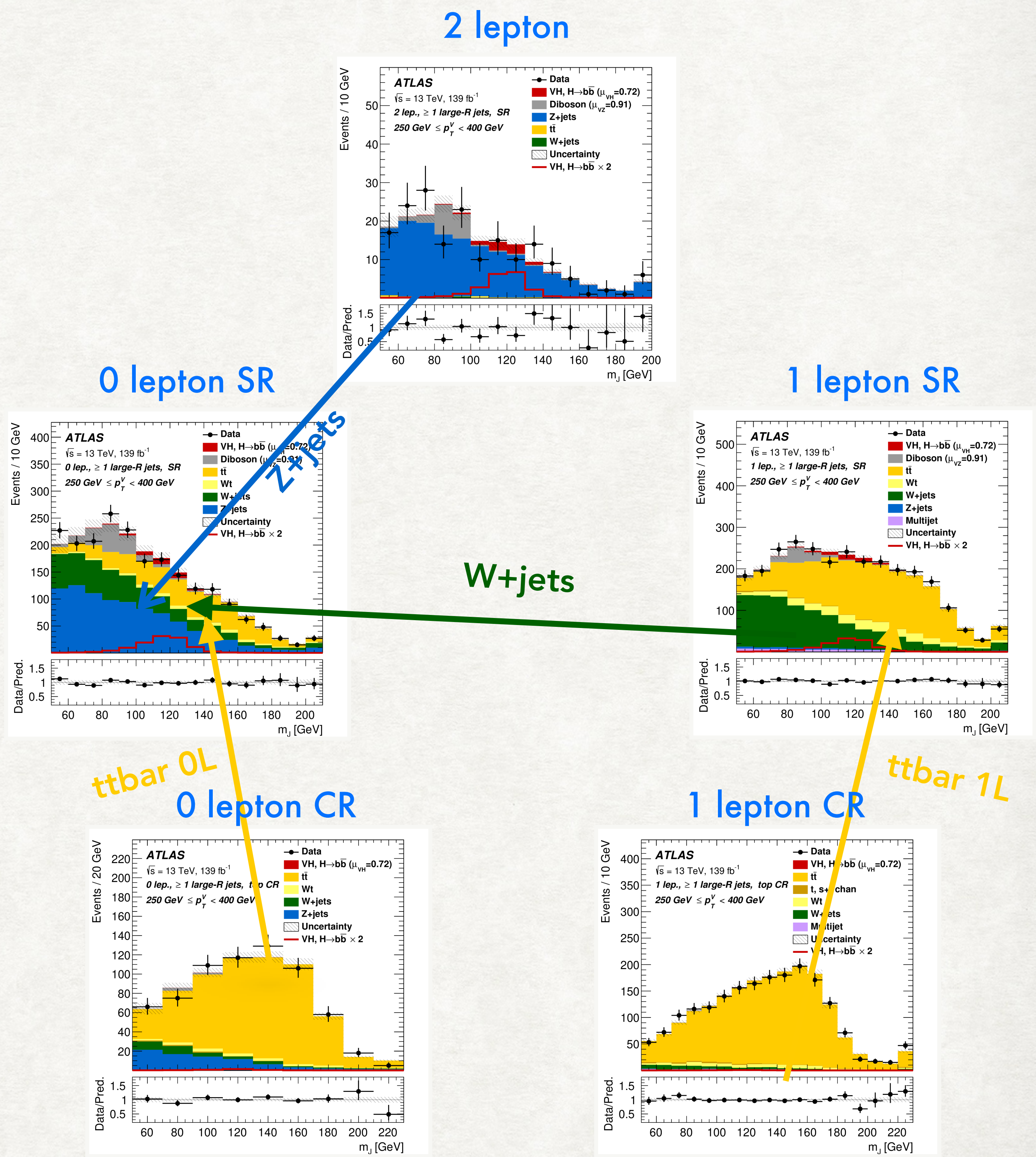




- Improve mass resolution using
  - Muon-in-jet correction
  - Kinematic fit (2 lepton channel)

Background modelling

- Use state-of-the-art MC generators
- Main backgrounds have floating normalizations and are estimated in the fit (ttbar/V+jets)
- Other normalizations are estimated from MC (single top/VV)
- Shape uncertainties on m_J
- Parametrize extrapolation uncertainties across regions



Systematic Uncertainties

There are many source of systematics:

- Experimental uncertainties

  Large-R jets, small-R jets, MET, leptons, b-tagging, pile-up, luminosity…

- Simulated sample uncertainties

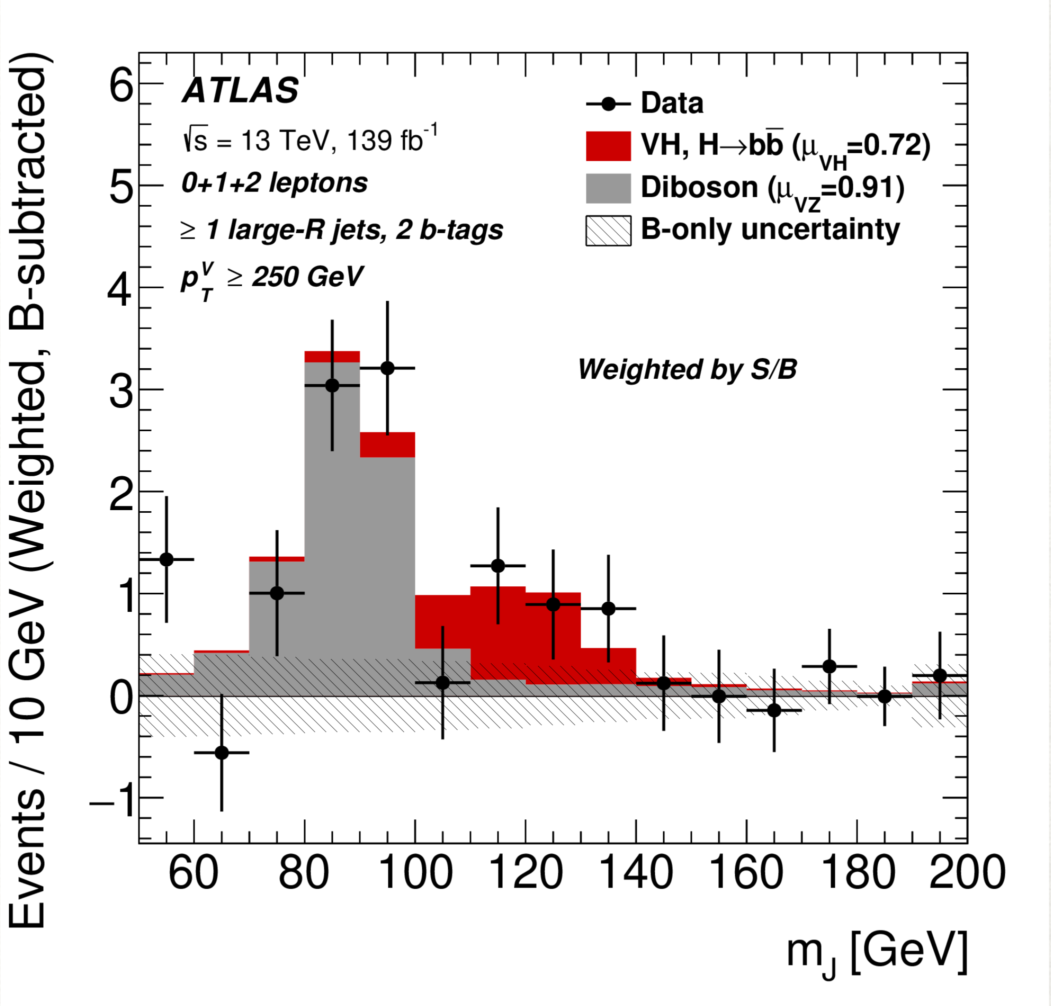
  Normalization, acceptance, m_J shape and MC stat. uncertainties…

- Multi-jet background uncertainties (1 lepton channel)

These uncertainties are taken into account in the final fitting model

Statistical Analysis

- Binned likelihood fitting has been performed to extract signal strength and significance.
- The impact of systematics are considered as nuisance parameters
- 14 categories are used in the combined fitting
- The likelihood function build on the m_J distributions



m_J

Signal strength

Nuisance Parameter

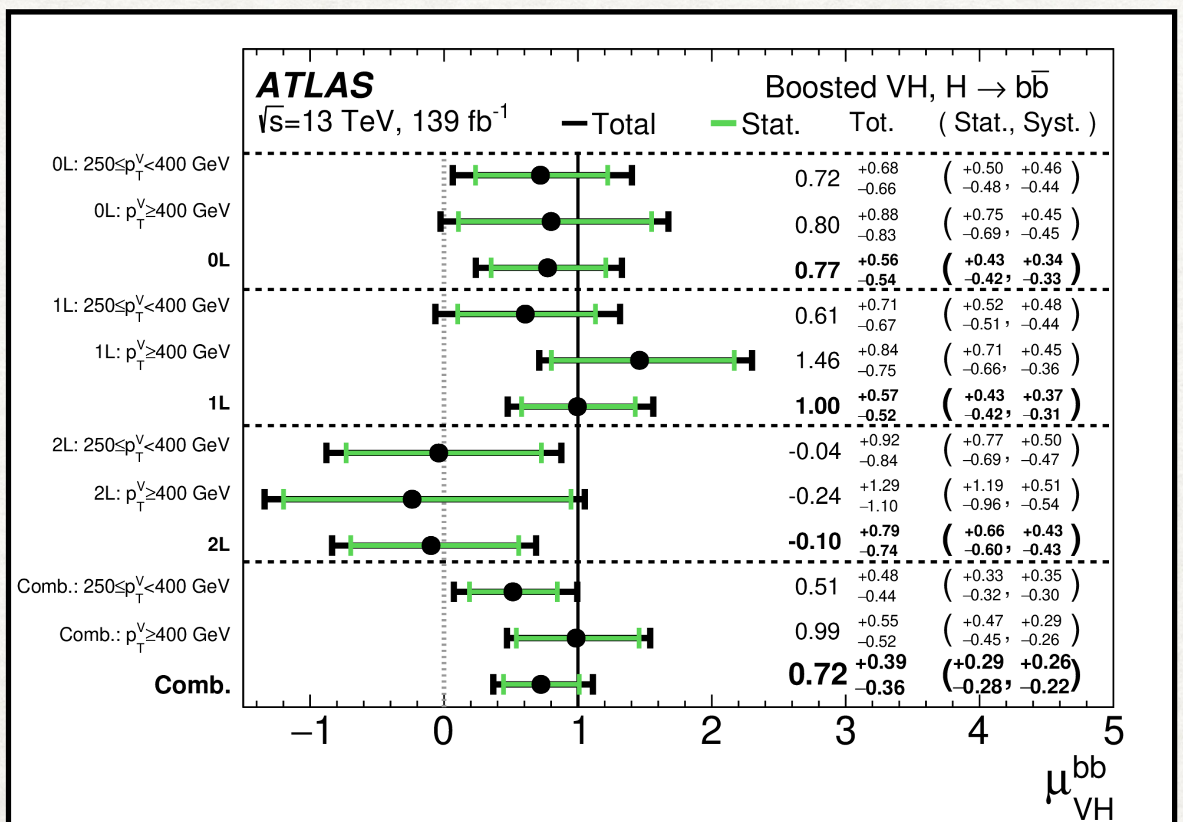
$$L = \left(\prod_{b \in bins} Pois(n_b | \mu S_b + B_b) \right) \left(\prod_{\theta \in NPs} Gauss(\theta, \sigma_\theta) \right)$$

Likelihood function

Results

VH(bb) boosted analysis

- Significance at **2.1σ (2.7σ exp.)**
- $\mu_{VHbb}=0.72\pm0.37$ compatible with SM prediction
- Analysis **statistically** limited
- Many important sources of uncertainties
  - Large-R jets, signal modeling, background modeling and MC stats



Signal strength

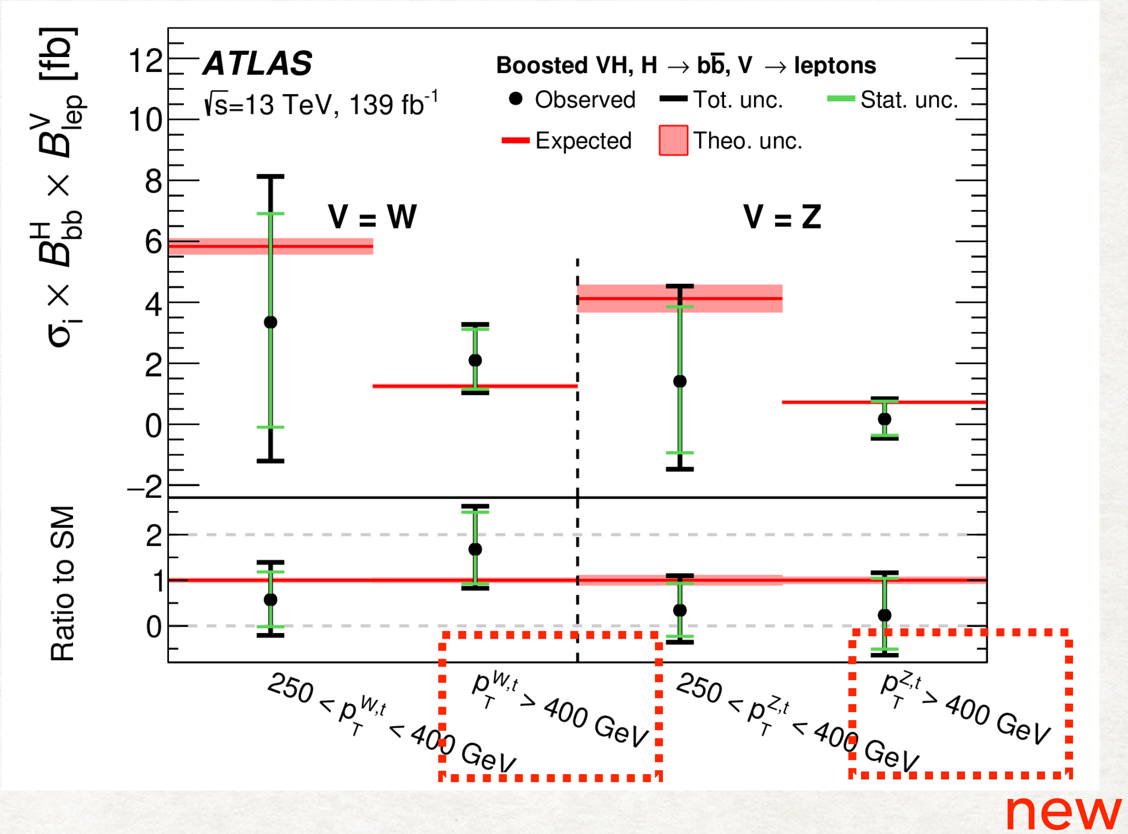
Source of uncertainty		Avg. impact
Total		0.372
Statistical		0.283
Systematic		0.240
Experimental uncertainties		
Small-R jets		0.038
Large-R jets		0.133
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.007
Leptons		0.010
b-tagging	b-jets	0.016
	c-jets	0.011
	light-flavour jets	0.008
	extrapolation	0.004
Pile-up		0.001
Luminosity		0.013
Theoretical and modelling uncertainties		
Signal		0.038
Backgrounds		0.100
↪ Z + jets		0.048
↪ W + jets		0.058
↪ $t\bar{t}$		0.035
↪ Single top quark		0.027
↪ Diboson		0.032
↪ Multijet		0.009
MC statistical		0.092

Breakdown of uncertainties

Results

Simplified Template Cross Section (STXS)

- Differential cross section measurement in "simplified" fiducial regions
- Sensitive to potential BSM physics
- STXS measurement in 2 p_T^V (GeV) bins: [250,400] and [400,∞]

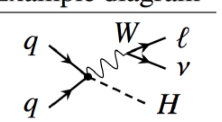
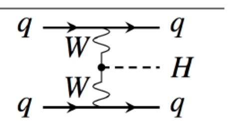
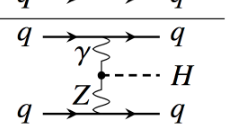
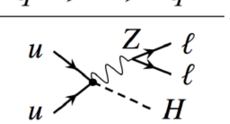
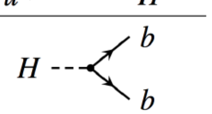


STXS region ($\|y_H\| < 2.5, H \to b\bar{b}$)	SM prediction [fb]	Result	(Tot.)	(Stat.)	(Syst.) [fb]
$W \to \ell\nu;\ p_T^{W,t} \in [250, 400]$ GeV	5.83 ± 0.26	3.3	$^{+4.8}_{-4.6}$	$^{+3.6}_{-3.4}$	$^{+3.2}_{-3.0}$
$W \to \ell\nu;\ p_T^{W,t} \in [400, \infty]$ GeV	1.25 ± 0.06	2.1	$^{+1.2}_{-1.1}$	$^{+1.0}_{-0.9}$	$^{+0.6}_{-0.5}$
$Z \to \ell\ell, \nu\nu;\ p_T^{Z,t} \in [250, 400]$ GeV	4.12 ± 0.45	1.4	$^{+3.1}_{-2.9}$	$^{+2.4}_{-2.3}$	$^{+1.9}_{-1.7}$
$Z \to \ell\ell, \nu\nu;\ p_T^{Z,t} \in [400, \infty]$ GeV	0.72 ± 0.05	0.2	$^{+0.7}_{-0.6}$	$^{+0.6}_{-0.5}$	$^{+0.3}_{-0.3}$

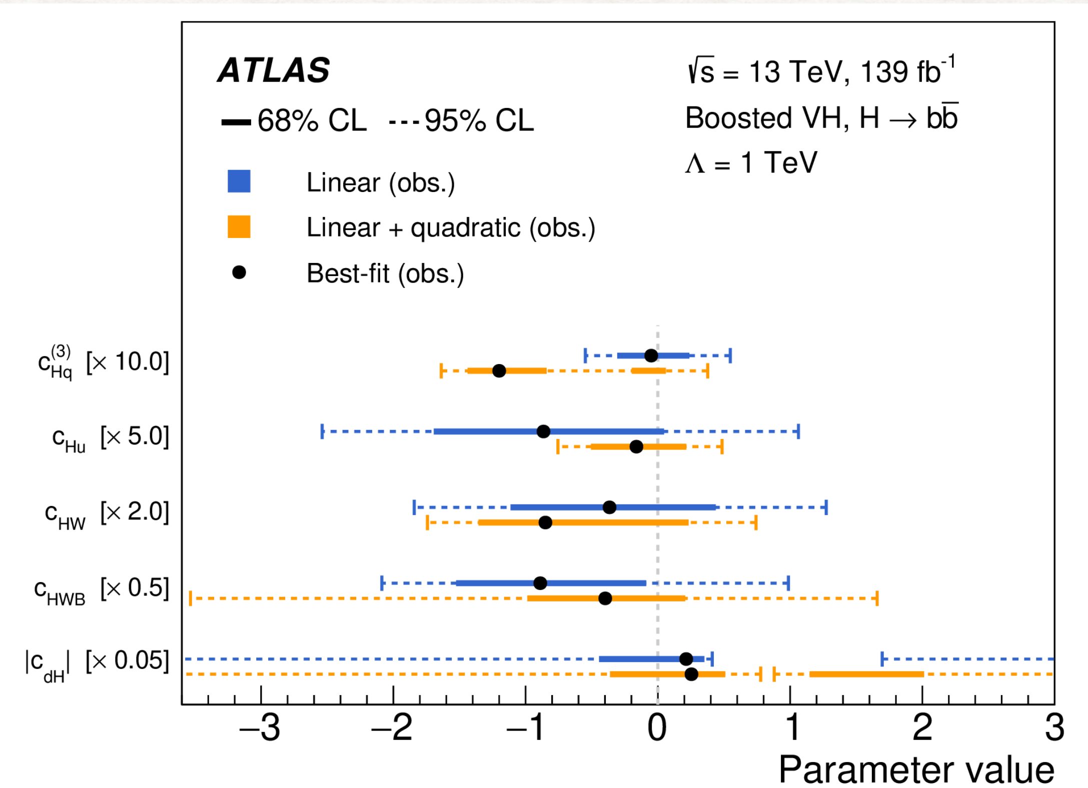
Results

EFT interpretation of STXS results

- Interpret STXS results using and Effective Field Theory (EFT) approach
- Model independent way to probe BSM physics
- STXS results parameterized as linear/quadratic polynomials in c_i coefficients
- Extract 68%/95% CL limits for leading c_i coefficients

Wilson coefficient	Operator definition	Example diagram	
c_{Hq3}	$(i\Phi^\dagger \overleftrightarrow{D}^I_\mu \Phi)(\bar{q}\sigma^I\gamma^\mu q)$		→ p_T^V dependent, affects WH vertex
c_{HW}	$\Phi^\dagger\Phi W^I_{\mu\nu}W^{I\mu\nu}$		→ affects HWW vertex
c_{HWB}	$\Phi^\dagger\Phi W^I_{\mu\nu}B^{I\mu\nu}$		→ affects HZγ vertex
c_{Hu}	$(i\Phi^\dagger \overleftrightarrow{D}^I_\mu \Phi)(\bar{u}\gamma^\mu u)$		→ affects ZH vertex
$\lvert c_{dH}\rvert$	$(\Phi^\dagger\Phi)(\bar{q}d\Phi)$		→ affects Hbb vertex directly

Definitions of Leading coefficients



Observed confidence interval @ 68% and 95% CL

Conclusion

- First VH(bb) analysis at high-p_T, using boosted techniques
  - Significance at 2.1**σ** (2.7**σ** exp.)
  - μ_{VHbb}=0.72±0.37 compatible with SM prediction
- Performed STXS measurement in an additional p_T^V bins
- EFT interpretation of the results
  - Plan to make it public in HEPdata

Backup

VZ analysis

Additional validation of the robustness

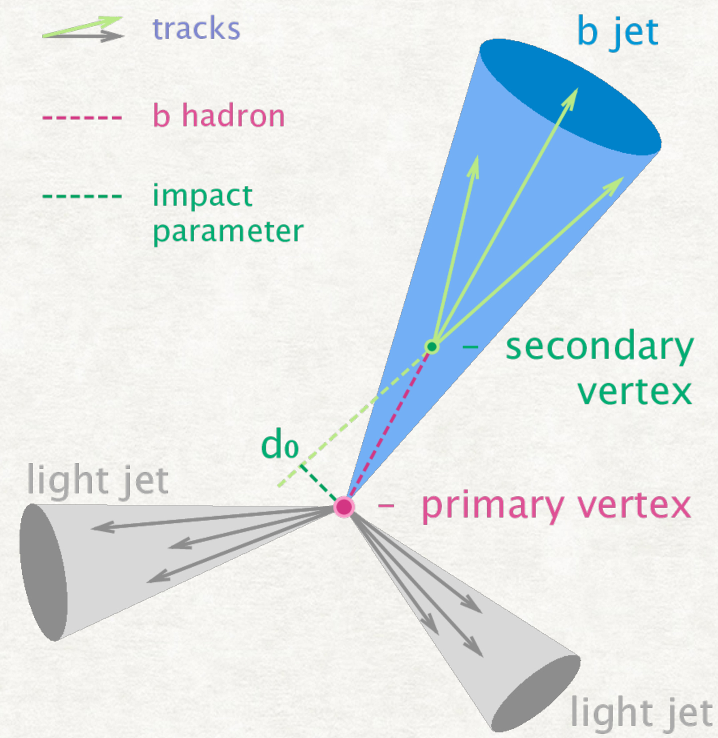
- Robust validation of background model and associated uncertainties
- Same analysis strategy as VHbb to search for VZ signal
  - 0 lepton: Z→νν, Z→bb
  - 1 lepton: W→lν, Z→bb
  - 2 lepton: Z→ll, Z→bb
- Results compatible with SM prediction

$$\mu_{VZ}^{bb} = 0.91^{+0.29}_{-0.23} = 0.91 \pm 0.15(\mathrm{stat.})^{+0.25}_{-0.17}(\mathrm{syst.}) \qquad \sigma_{VZ}^{bb} = 5.4\,(5.7)\;obs.\,(exp.)$$

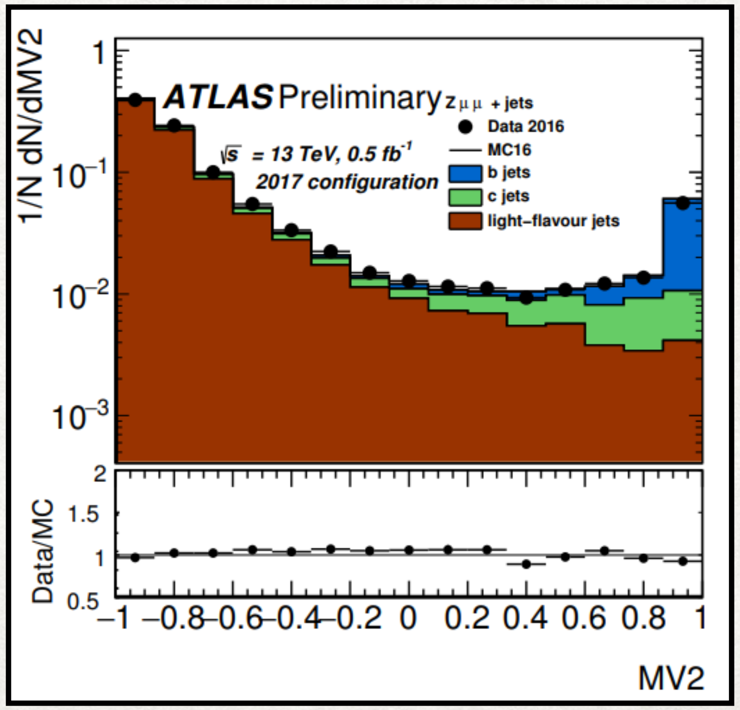
B-tagging

To identify jets from b-quark

- B-tagging
  - Depends on the good operation of the tracker
  - Performance in Run 2 relying on
    - New IBL detector installed in LS1 (2013-2014)
    - Tracking optimized for high-PU and high-p_T
    - Better ML algorithm



- MV2C10 tagger used in this analysis
  - Boosted Decision Tree (BDT) algorithm
  - Exploits the features of b-jets
    - Secondary vertex (SV)
    - Impact parameter (IP)
    - B-hadron decay chain inside jet code (JetFitter)



Rejection of light / c jets 300 / 8 at 70% b-jets efficiency

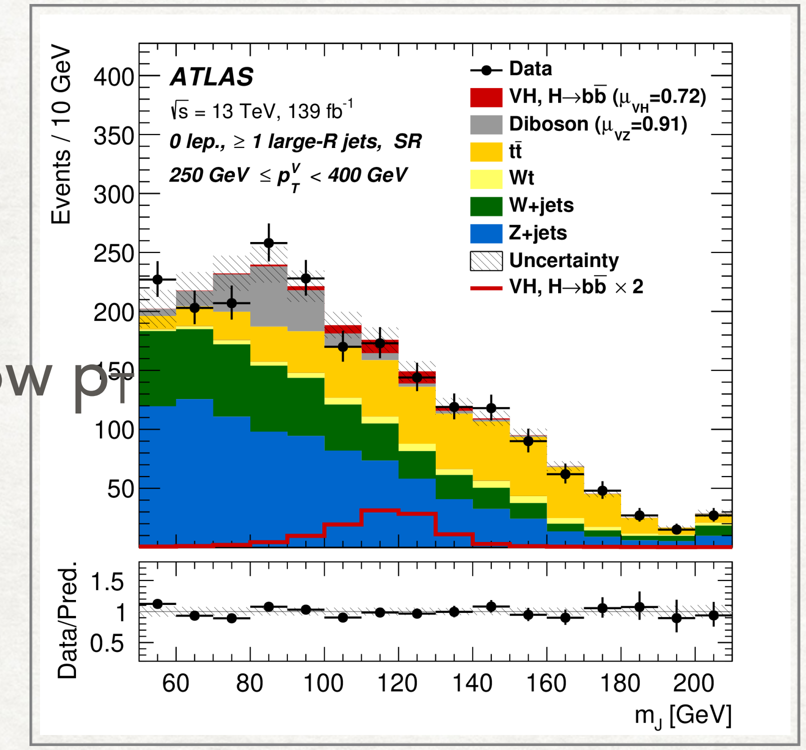
Simulation samples

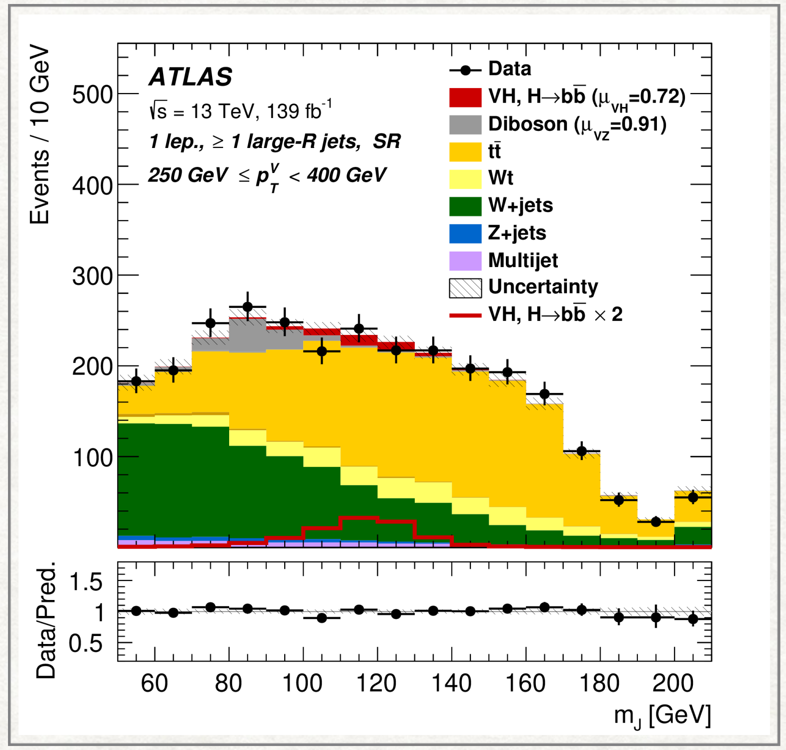
Process	ME generator	ME PDF	PS and Hadronisation	UE model tune	Cross-section order
Signal ($m_H = 125$ GeV and $b\bar{b}$ branching fraction set to 58%)					
$qq \to WH \to \ell\nu b\bar{b}$	Powheg-Box v2 [34] + GoSam [36] + MiNLO [37,38]	NNPDF3.0NLO$^{(\star)}$ [30]	Pythia 8.212 [35]	AZNLO [23]	NNLO(QCD)+ NLO(EW) [39,40,41,42,43,44,45]
$qq \to ZH \to \nu\nu b\bar{b}/\ell\ell b\bar{b}$	Powheg-Box v2 + GoSam + MiNLO	NNPDF3.0NLO$^{(\star)}$	Pythia 8.212	AZNLO	NNLO(QCD)$^{(\dagger)}$+ NLO(EW)
$gg \to ZH \to \nu\nu b\bar{b}/\ell\ell b\bar{b}$	Powheg-Box v2	NNPDF3.0NLO$^{(\star)}$	Pythia 8.212	AZNLO	NLO+ NLL [46,47,48,49,50]
Top quark ($m_t = 172.5$ GeV)					
$t\bar{t}$	Powheg-Box v2 [34,51]	NNPDF3.0NLO	Pythia 8.230	A14 [24]	NNLO+NNLL [52]
s-channel	Powheg-Box v2 [53,34]	NNPDF3.0NLO	Pythia 8.230	A14	NLO [54]
t-channel	Powheg-Box v2 [53,34]	NNPDF3.0NLO	Pythia 8.230	A14	NLO [55]
Wt	Powheg-Box v2 [56,34]	NNPDF3.0NLO	Pythia 8.230	A14	Approximate NNLO [57]
Vector boson + jets					
$W \to \ell\nu$	Sherpa 2.2.1 [25,26,27,28]	NNPDF3.0NNLO	Sherpa 2.2.1 [58,59]	Default	NNLO [60]
$Z/\gamma^* \to \ell\ell$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO
$Z \to \nu\nu$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO
Diboson					
$qq \to WW$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$qq \to WZ$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$qq \to ZZ$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$gg \to VV$	Sherpa 2.2.2	NNPDF3.0NNLO	Sherpa 2.2.2	Default	NLO

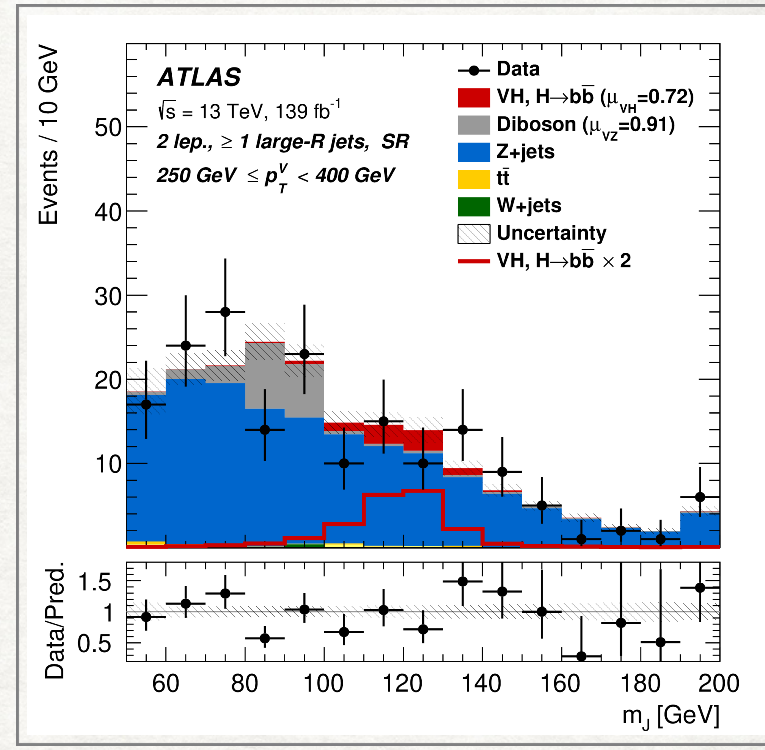
Post-fit m$_J$ distributions in SR

0L 1L 2L

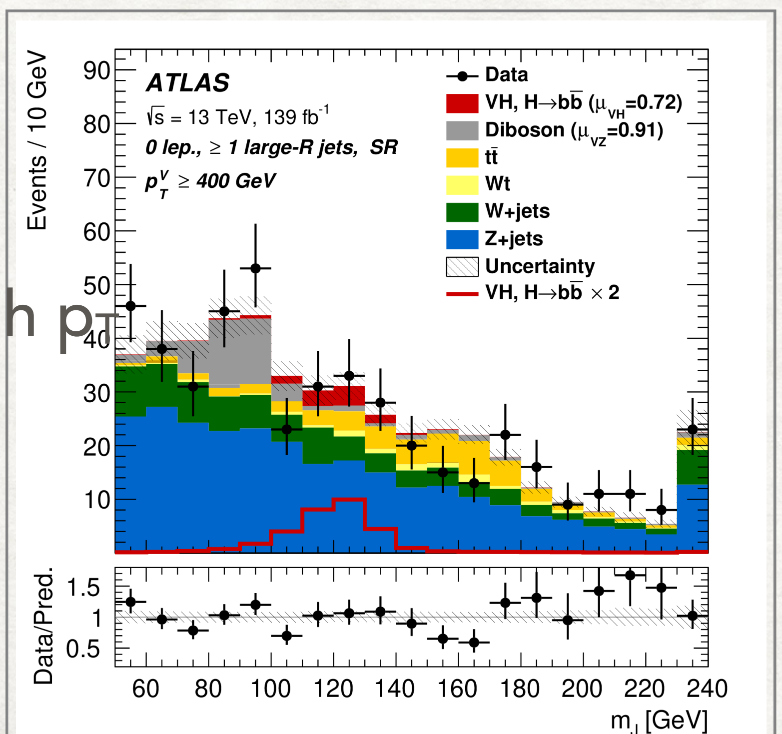
Low p$_T$

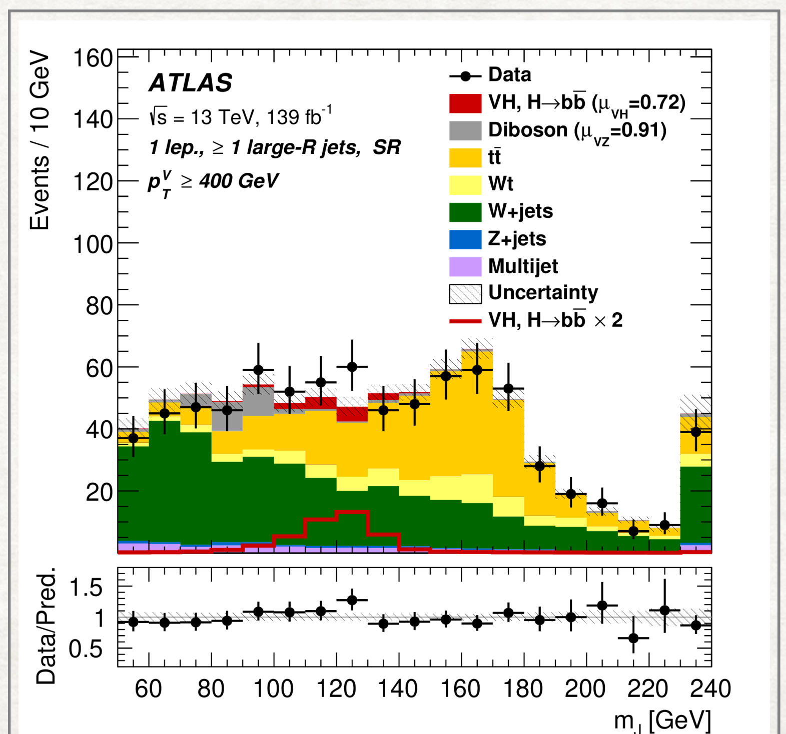


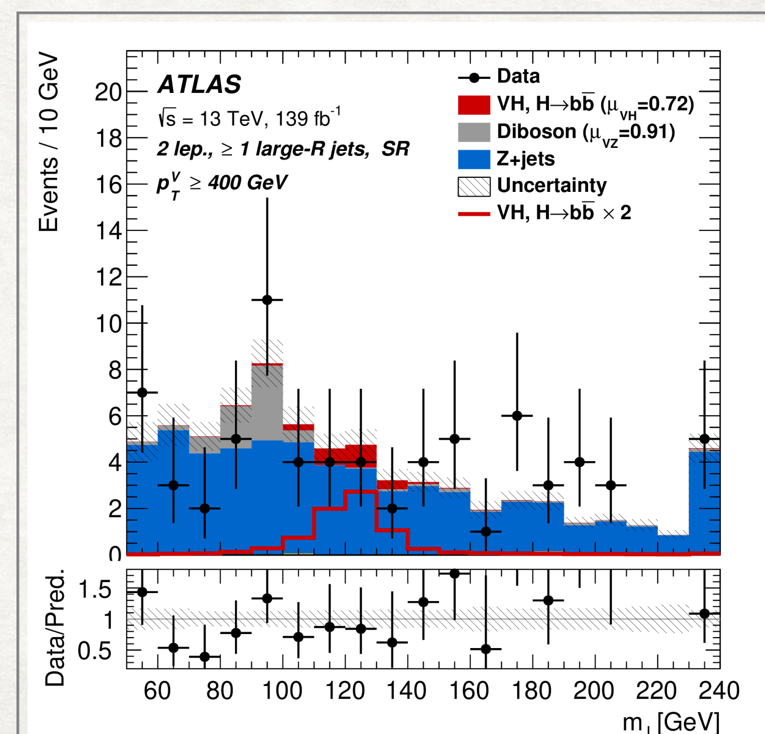




High p$_T$





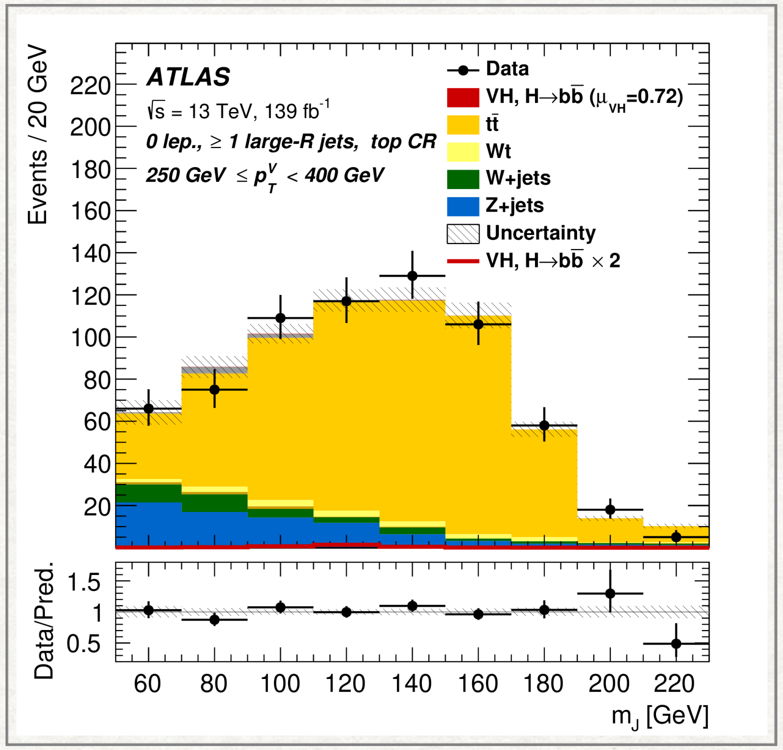


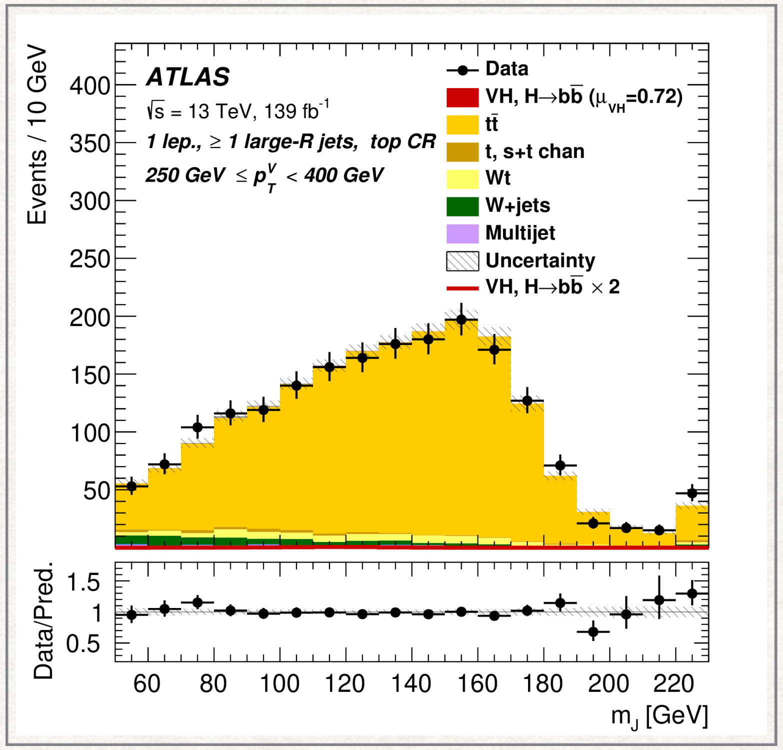
Post-fit m_J distributions in top CR

0L

1L

Low p_T





High p_T