Studies of Higgs boson properties in decays to two b-quarks at the ATLAS experiment

Weitao Wang
on behalf of the ATLAS collaboration

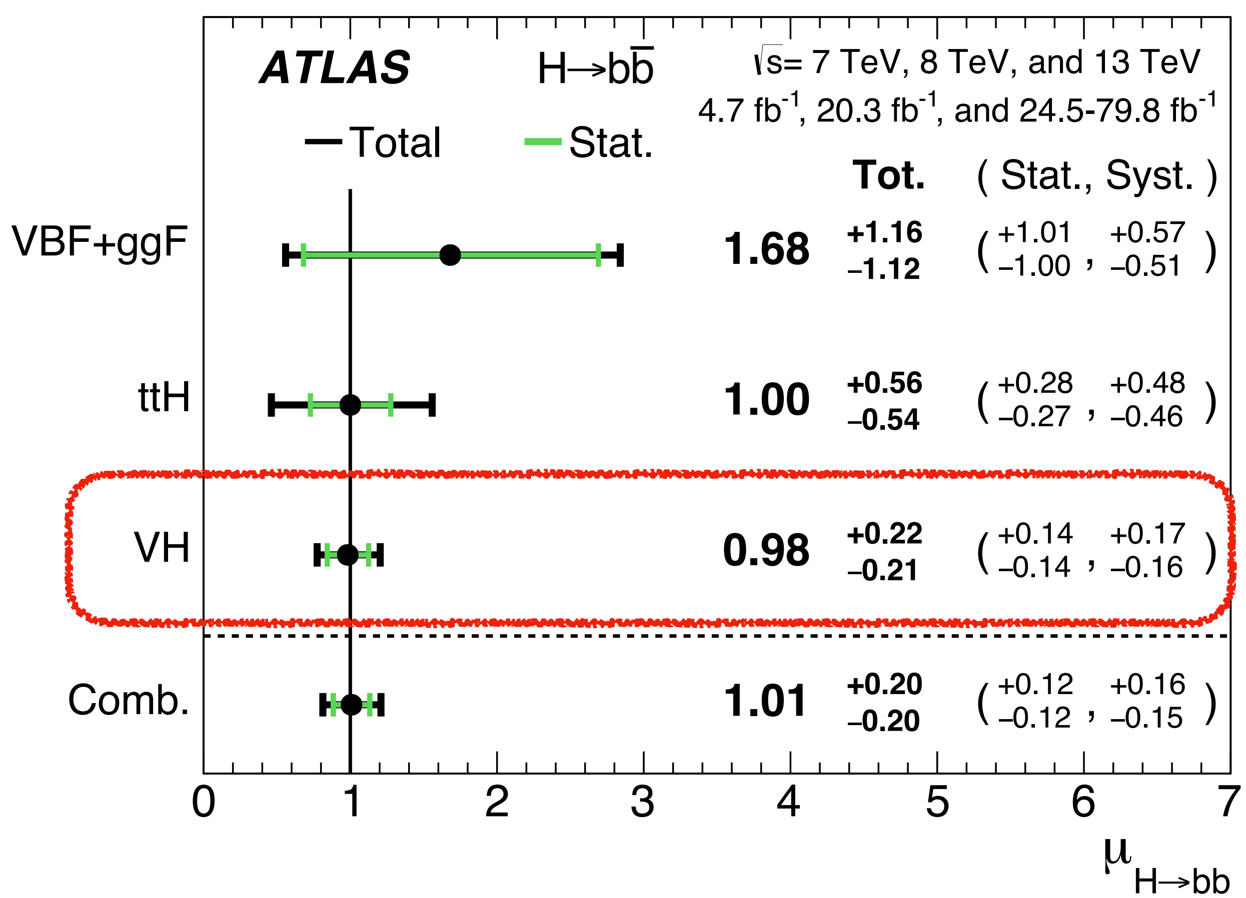
Phenomenology 2020 Symposium
05.05.2020

Higgs Boson cross-section measurements at ATLAS

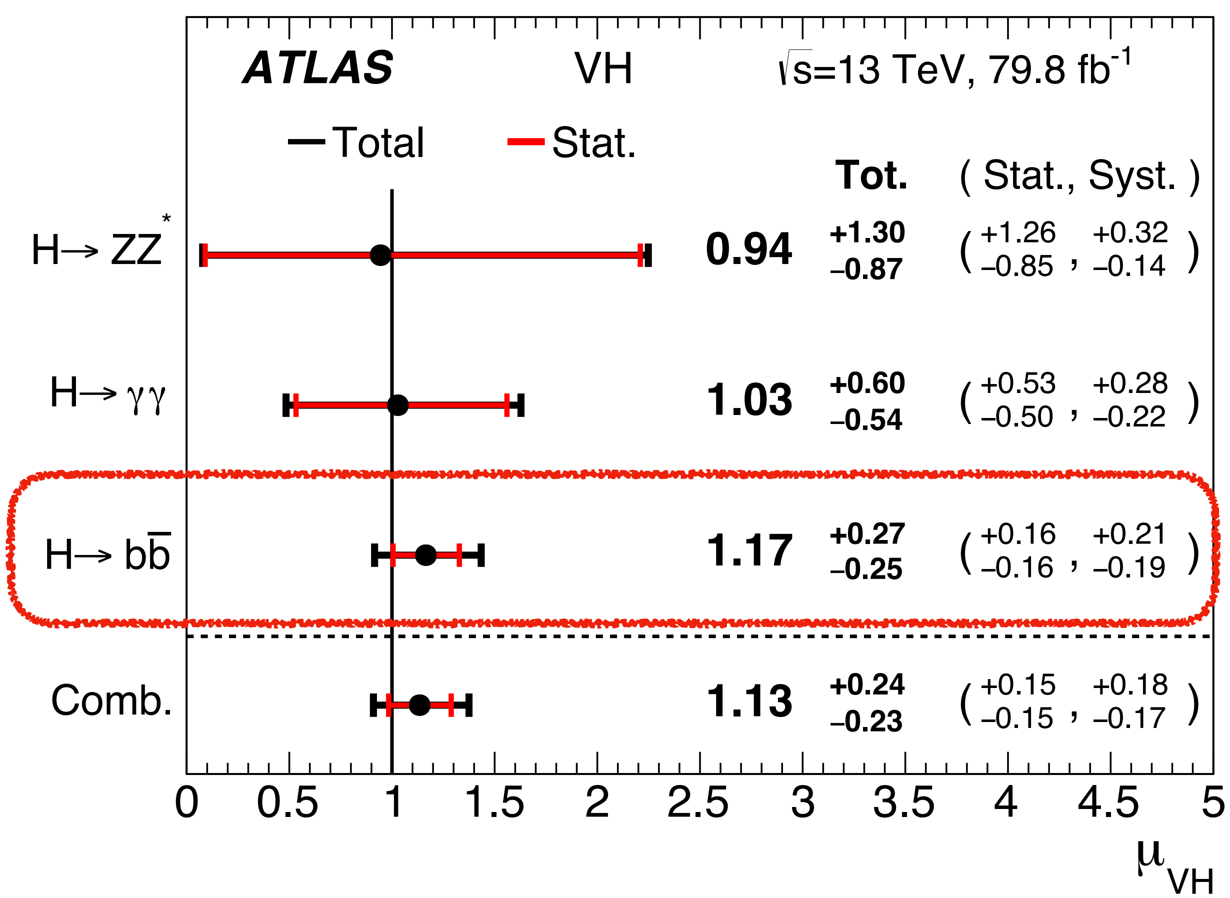
$H \to b\bar{b}$ is the dominant decay mode of Higgs boson — branch fraction **58%**

Cross-section of $H \to b\bar{b}$ decay has been measured at ATLAS in *VBF*, *VH* and $t\bar{t}H$ processes

$H \to b\bar{b}$ signal strength Run-1 + Run-2



VH signal strength Run-2



Phys. Lett. B 786 (2018) 59

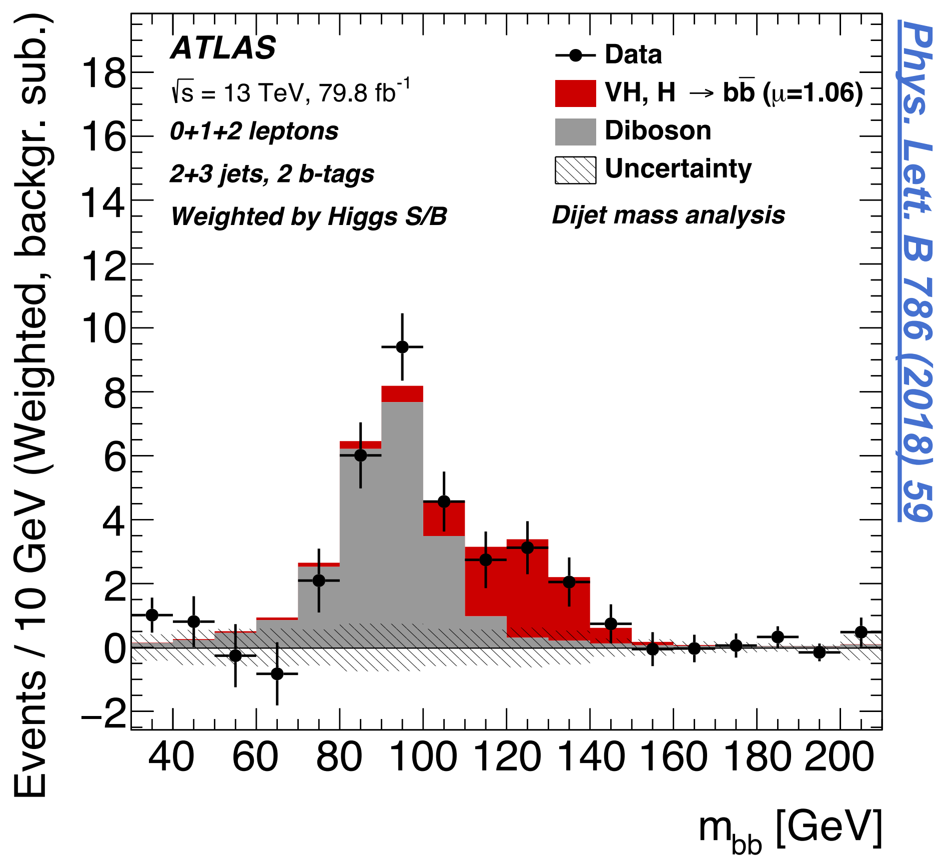
VH(→bb) process presents the best results in both $H \to b\bar{b}$ and VH measurement

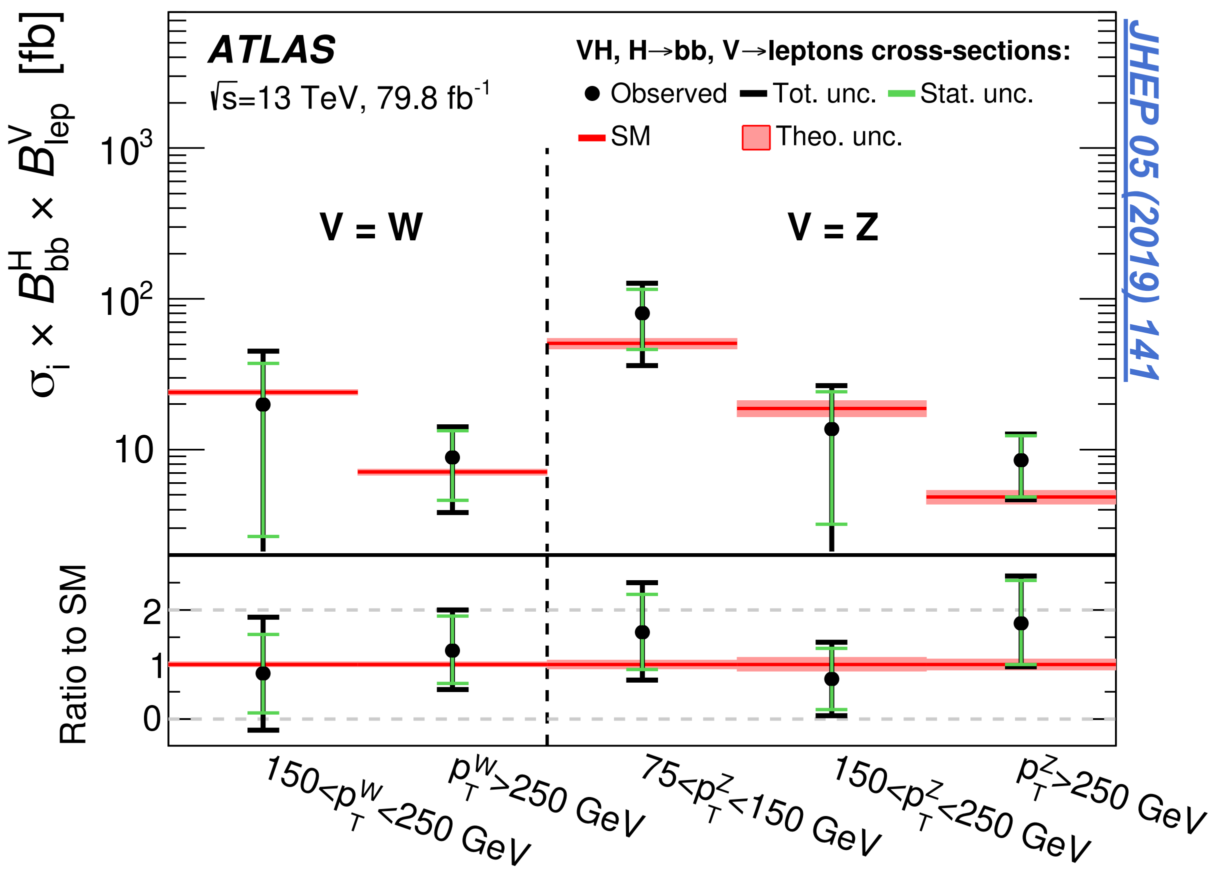
VH(→bb): the golden channel for H→bb and VH studies

VH(bb) analysis

Partial Run-2 79.8 fb^{-1} results

- 4.9σ significance on VH(bb) channel
- Observation of $H \rightarrow b\bar{b}$ in combination with other production modes
- Simplified template cross-section (STXS): in 2 p_T^W and 3 p_T^Z STXS bins





Aims for full Run-2 139fb^{-1} data: *two standalone analyses*

Resolved analysis

Improve the significance
and precision of cross-section measurement

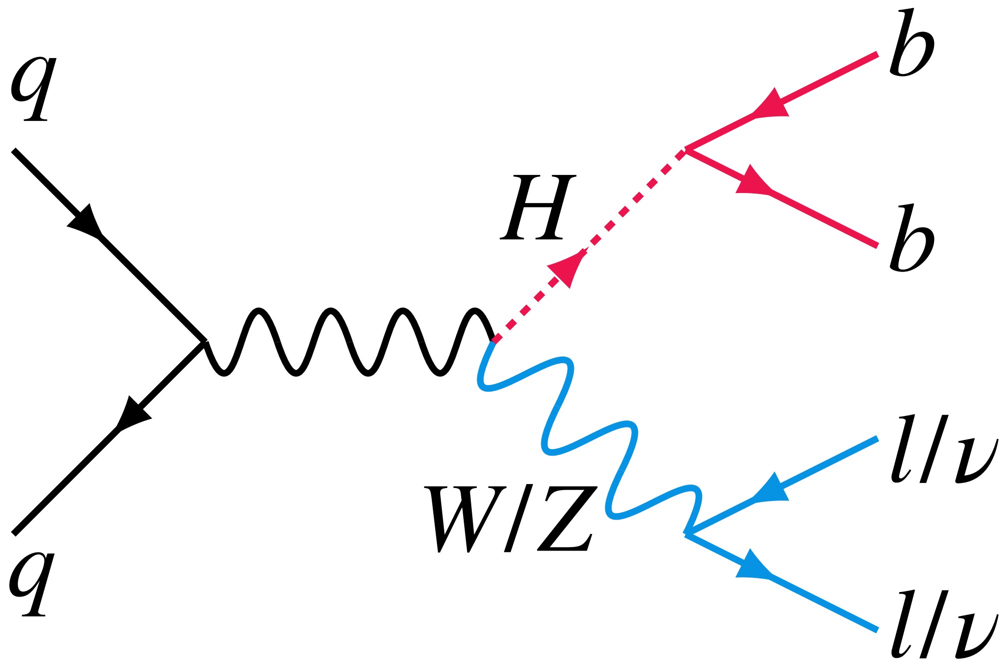
ATLAS-CONF-2020-006

Boosted analysis

New $H \rightarrow b\bar{b}$ reconstruction techniques
Extend STXS bins to high p_T^V (> 400 GeV)

ATLAS-CONF-2020-007

Resolved VH(bb): Events selection



Higgs candidate:

Exactly 2 b-tagged anti-k_t R=0.4 jets
b-jet energy corrections applied to improve m_{bb} resolution

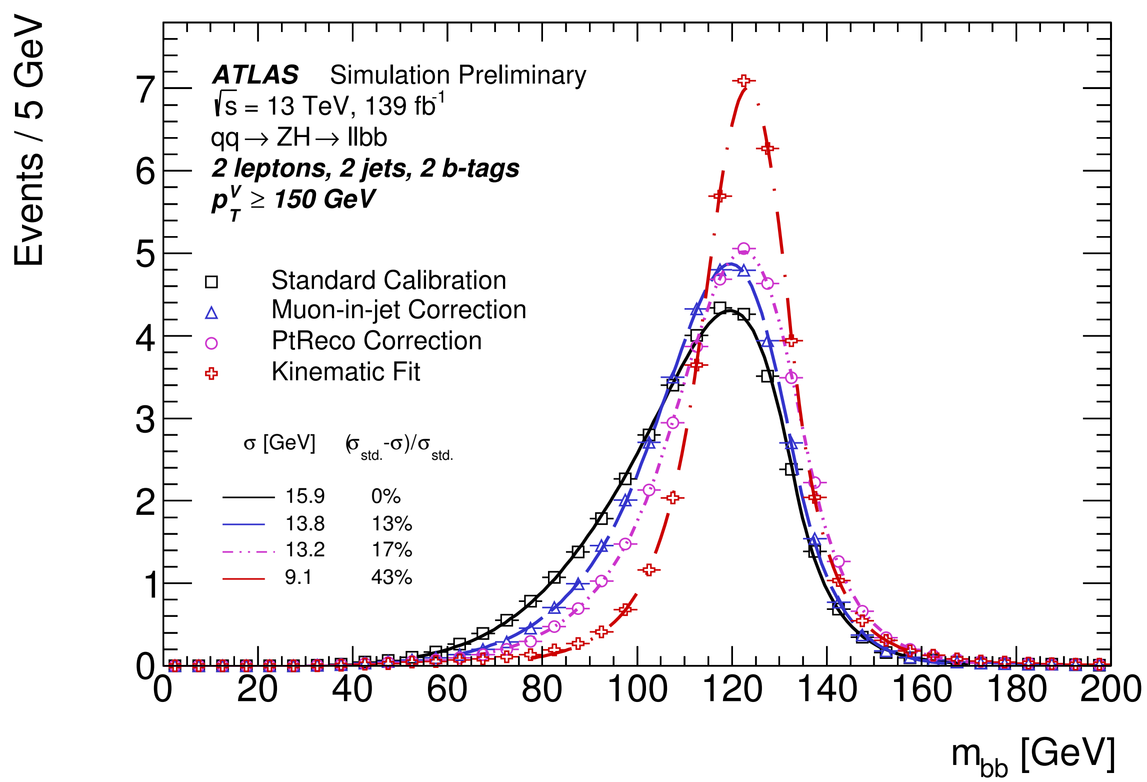
Vector boson candidate:

0L: $Z \rightarrow \nu\bar{\nu}$

1L: $W \rightarrow l\nu$

2L: $Z \rightarrow ll$

b-jet energy corrections:



0/1L: muon-in-jet and PtReco
2L: muon-in-jet and kinematic fit

Event Selection:

- 1 Higgs candidate
- 1 Vector Boson candidate
- Anti-multijet cut in 0/1L

Resolved VH(bb): Events categorisation

Event categorisation

Channel	0/1 Lepton	2 Leptons
Number of jets	2/3 jets	2/≥3 jets
p_T^V	150-250 GeV, >250 GeV	75-150GeV, 150-250 GeV, >250 GeV

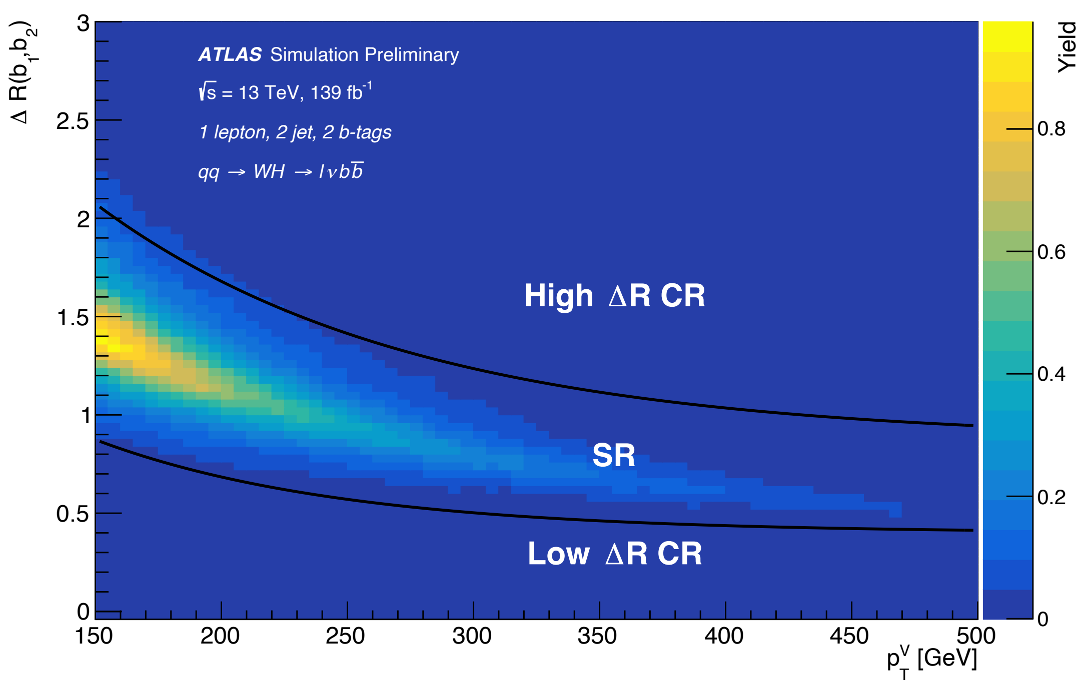
→ Improve significance

→ Corresponding to STXS bins

14 analysis categories

Control regions and signal regions

- Use jet-angular separation $\Delta R(b,b)$ to define control regions
- The control regions are used to constrain the normalisation of dominant background



High ΔR : $t\bar{t}$

SR: Contains > 80% signal events
(>93% in the 2-jet categories)

Low ΔR : $V + bb$

Total: 14 SR + 28 CR

Resolved VH(bb): Multivariate analysis

Boosted decision tree (BDT)

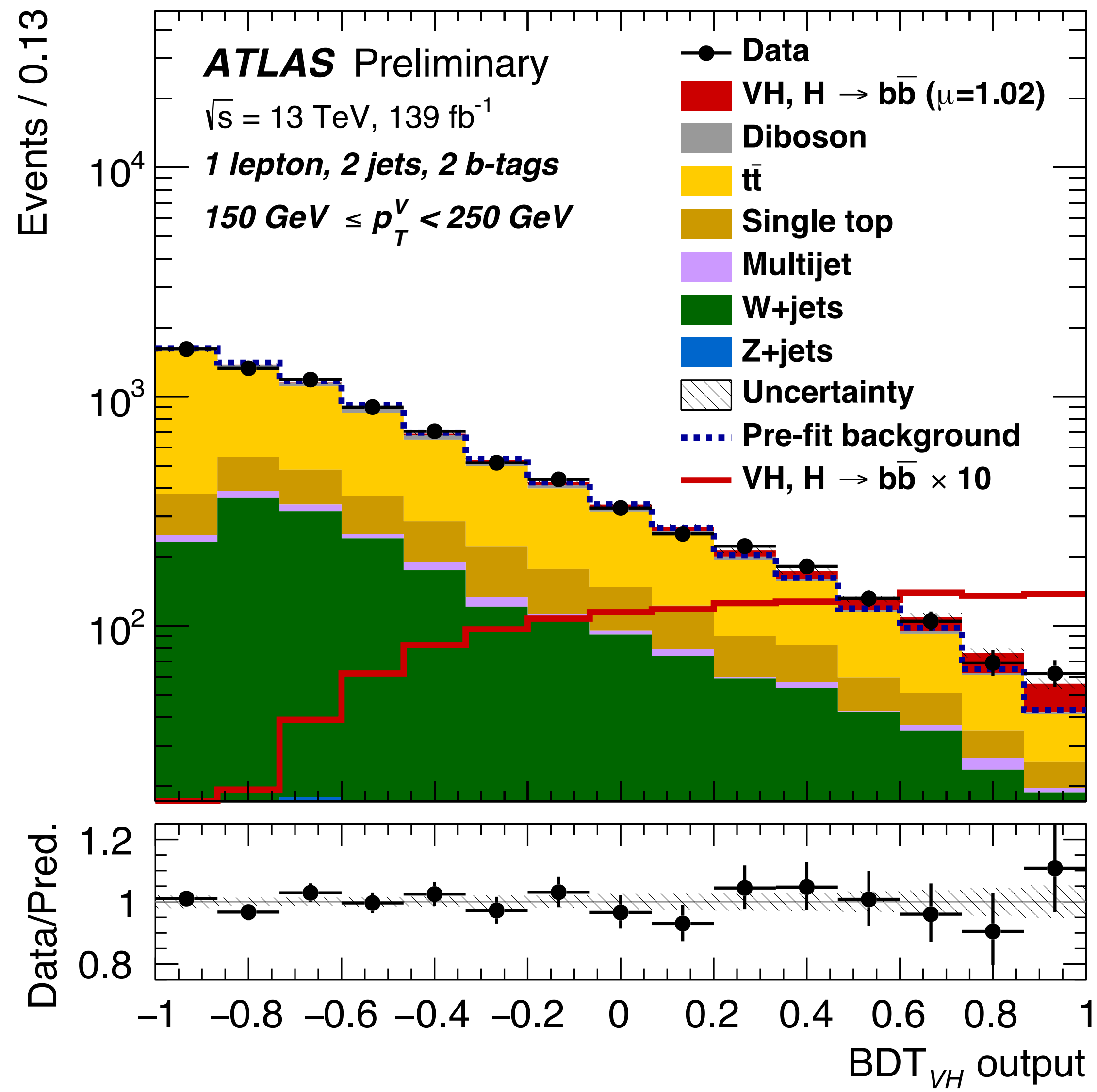
Variable	0-lepton	1-lepton	2-lepton
m_{bb}	×	×	×
$\Delta R(\vec{b_1}, \vec{b_2})$	×	×	×
$p_{\mathrm{T}}^{b_1}$	×	×	×
$p_{\mathrm{T}}^{b_2}$	×	×	×
p_{T}^{V}	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	×	×
$\Delta\phi(\vec{V}, \vec{bb})$	×	×	×
MV2(b_1)	×	×	
MV2(b_2)	×	×	
$\lvert\Delta\eta(\vec{b_1}, \vec{b_2})\rvert$	×		
m_{eff}	×		
$p_{\mathrm{T}}^{\mathrm{miss,st}}$	×		
$E_{\mathrm{T}}^{\mathrm{miss}}$	×	×	
$\min[\Delta\phi(\vec{\ell}, \vec{b})]$		×	
m_{T}^{W}		×	
$\lvert\Delta y(\vec{V}, \vec{bb})\rvert$		×	
m_{top}		×	
$\lvert\Delta\eta(\vec{V}, \vec{bb})\rvert$			×
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$			×
$m_{\ell\ell}$			×
$\cos\theta(\vec{\ell^-}, \vec{Z})$			×
Only in 3-jet events			
$p_{\mathrm{T}}^{\mathrm{jet}_3}$	×	×	×
m_{bbj}	×	×	×

Most powerful: m_{bb}, $\Delta R(\vec{b_1}, \vec{b_2})$, $p_{\mathrm{T}}^{b_1}$, $p_{\mathrm{T}}^{b_2}$

New: MV2(b_1), MV2(b_2); $p_{\mathrm{T}}^{\mathrm{miss,st}}$; $\cos\theta(\vec{\ell^-}, \vec{Z})$

Significance improvement:

- MV2(b_1), MV2(b_2): +7~10%
- $p_{\mathrm{T}}^{\mathrm{miss,st}}$: +2~3%
- $\cos\theta(\vec{\ell^-}, \vec{Z})$: +7%



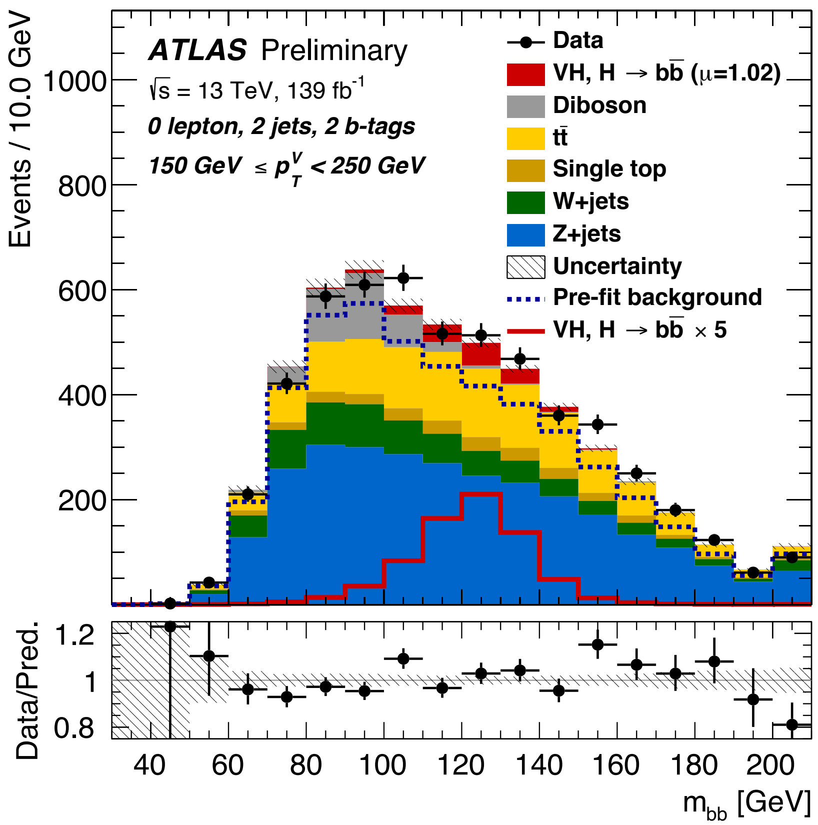
Distribution of BDT output in 1L

BDT output is the final discriminant

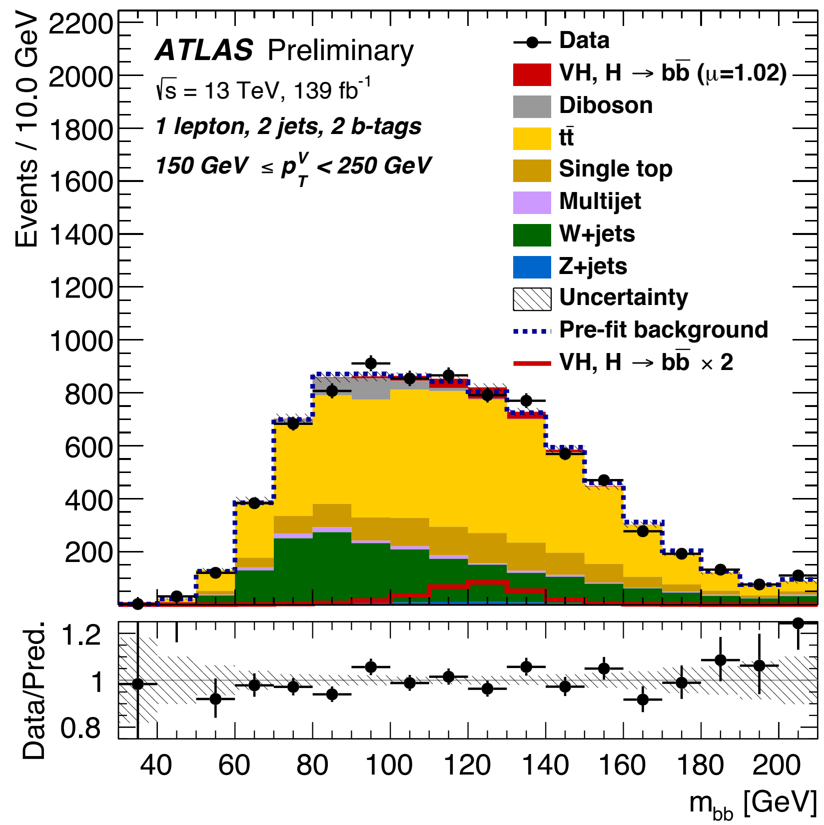
Resolved VH(bb): The fit model

Dominant backgrounds

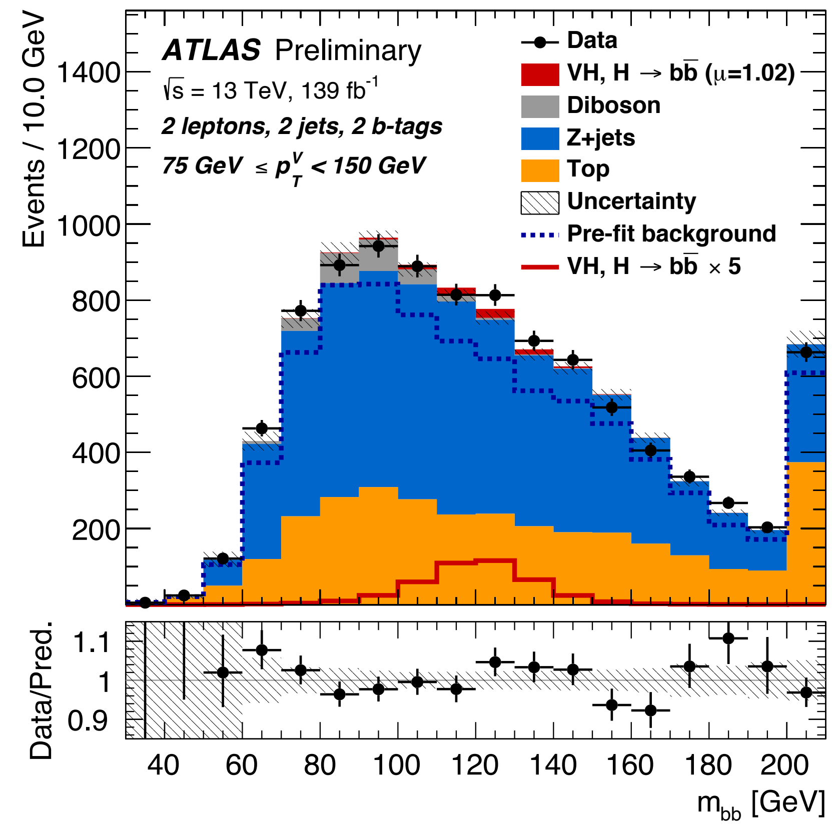
0L: W/Z+jets, $t\bar{t}$



1L: W+jets, $t\bar{t}$



2L: Z+jets, Top



The fit model

- Perform a binned likelihood fit simultaneously in the 14 signal regions and 28 control regions
- Normalisations of dominant backgrounds are **floated** in the final fit to data
- **Systematic uncertainties** included in form of nuisance parameters

Resolved VH(bb): Results

Observed (expected) significance:

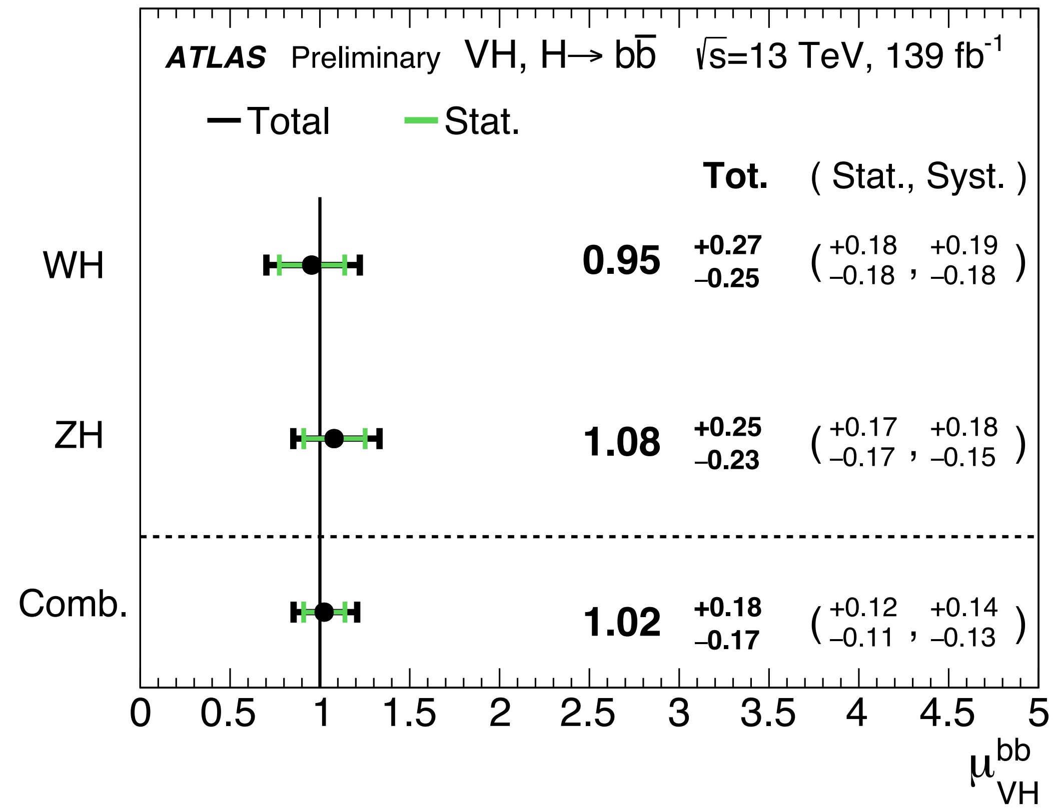
WH: 4.0 (4.1) σ
Strong evidence of WH production

ZH: 5.3 (5.1) σ
Observation of ZH production

VH: 6.7 (6.7) σ

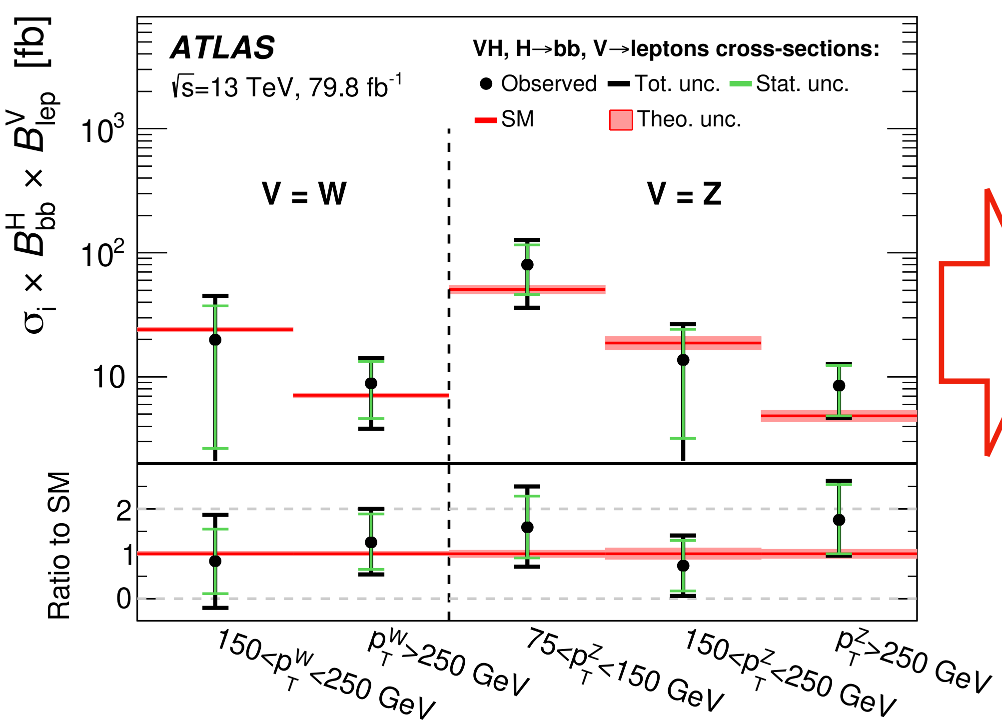
Statistical and systematic uncertainty of the same order

Signal strength

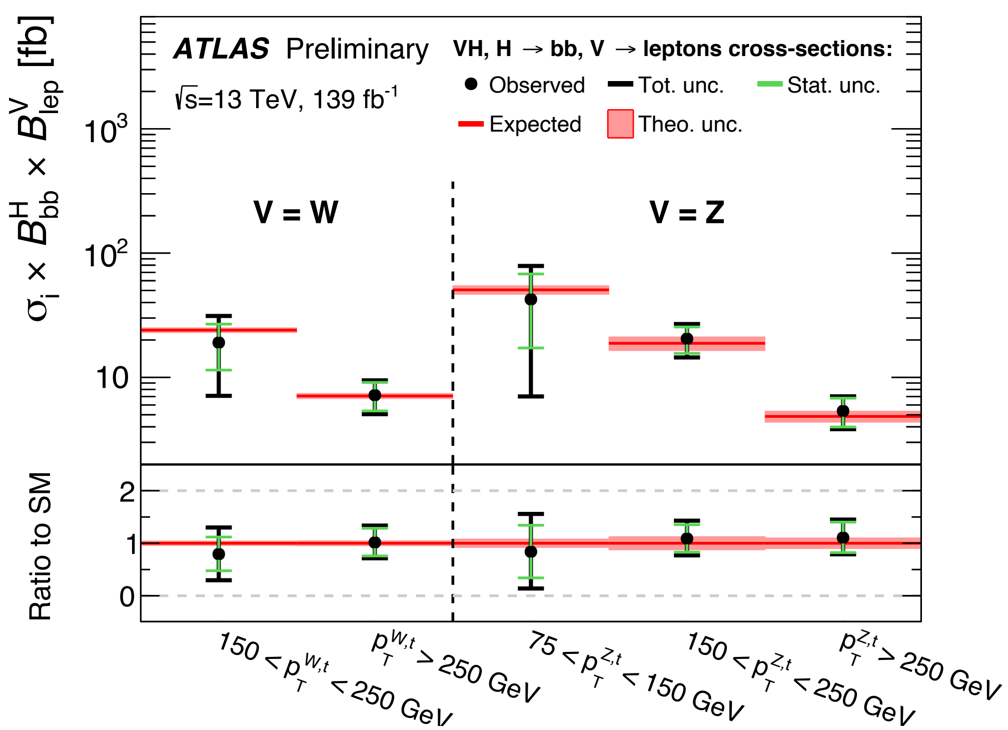


Cross-section measurement

Partial Run-2



Full Run-2



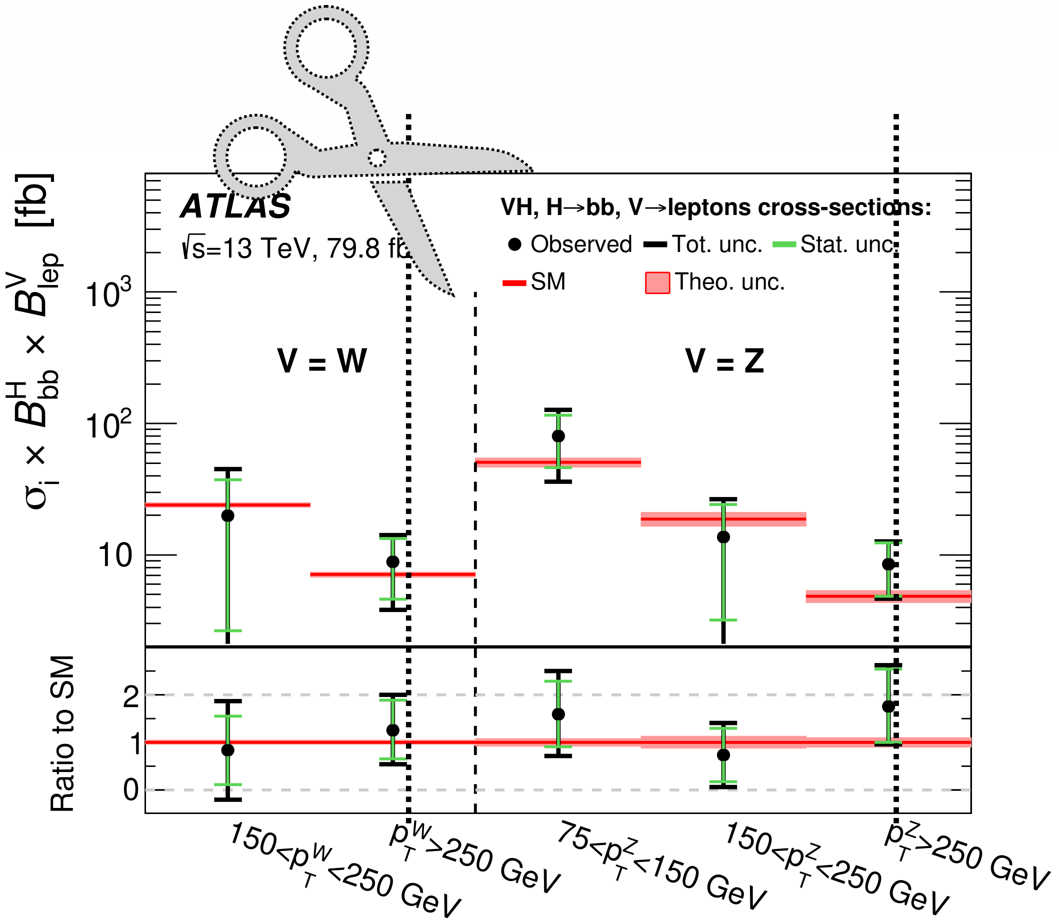
Significantly improve the precision of the Cross-section measurement
Introduce the $p_T^V > 250$ GeV analysis category:

- Strongly **reduce correlations** between STXS signal strengths and the uncertainty on high p_T^V bin

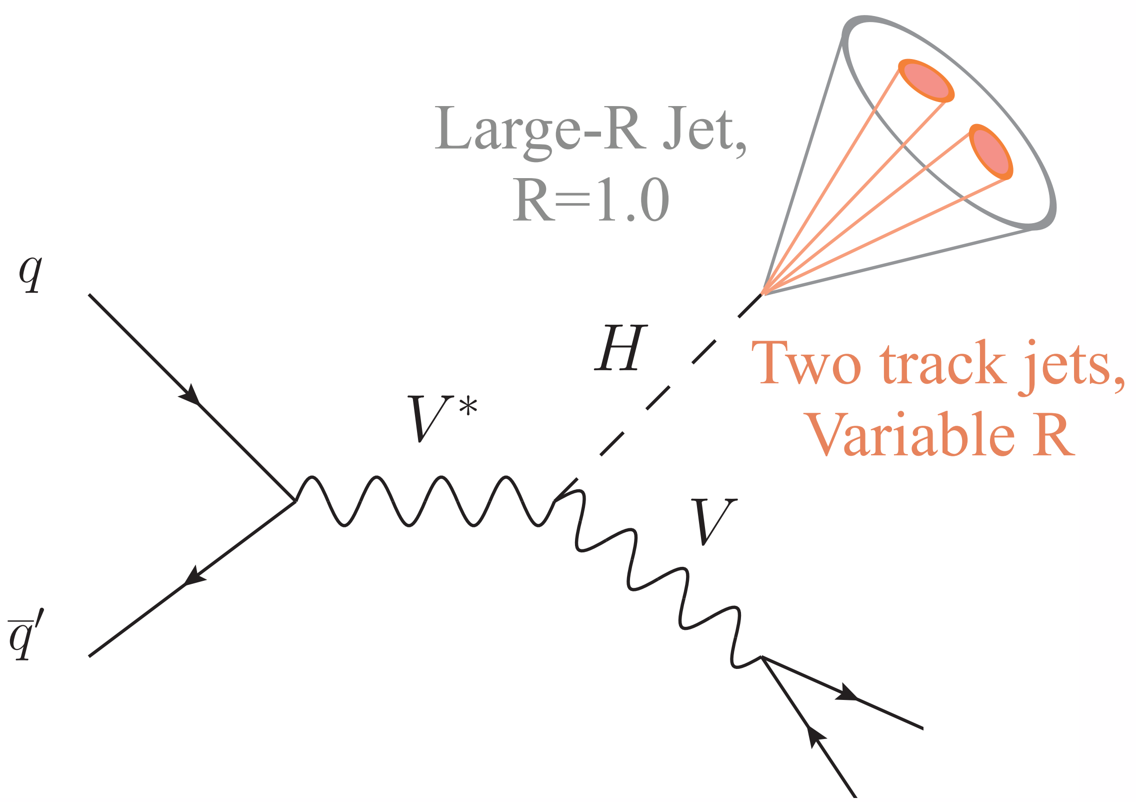
Approach to high pT — Boosted VH(bb) analysis

- Focus on $p_T^V > 250$ GeV region
- Extend the STXS bins:
  250 GeV $< p_T^Z <$ 400 GeV & $p_T^V >$ 400 GeV
- Establish boosted H(bb) reconstruction techniques for Higgs measurements

→ **Robust and standalone analysis**
Significant overlap with resolved VH(bb) analysis



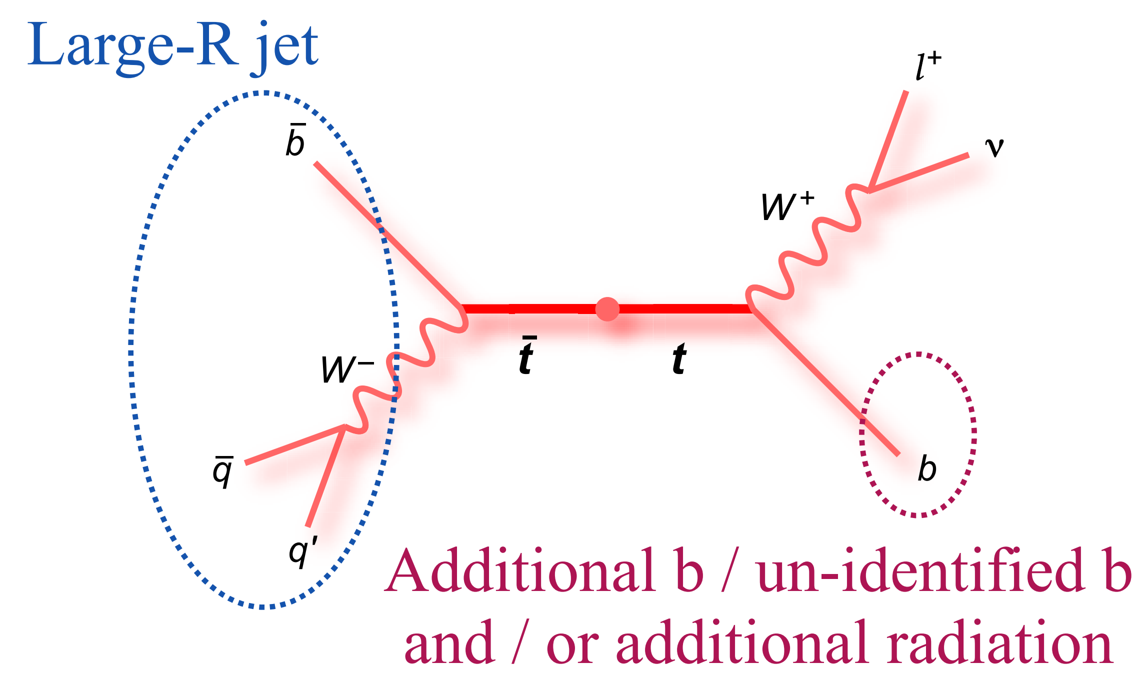
Boosted H→ bb reconstruction



- **Large-R jet**: LCTopo, R=1.0, trimmed
  $p_T > 250$ GeV, $|\eta| < 2.0$
- **Track jet:** variable radius
  $p_T > 10$ GeV, $|\eta| < 2.5$

- At least 2 track jets associated to large-R jet
- Leading two track jets in large-R jet are b-tagged (MV2c10@70%)

Boosted VH(bb): Event categorisation



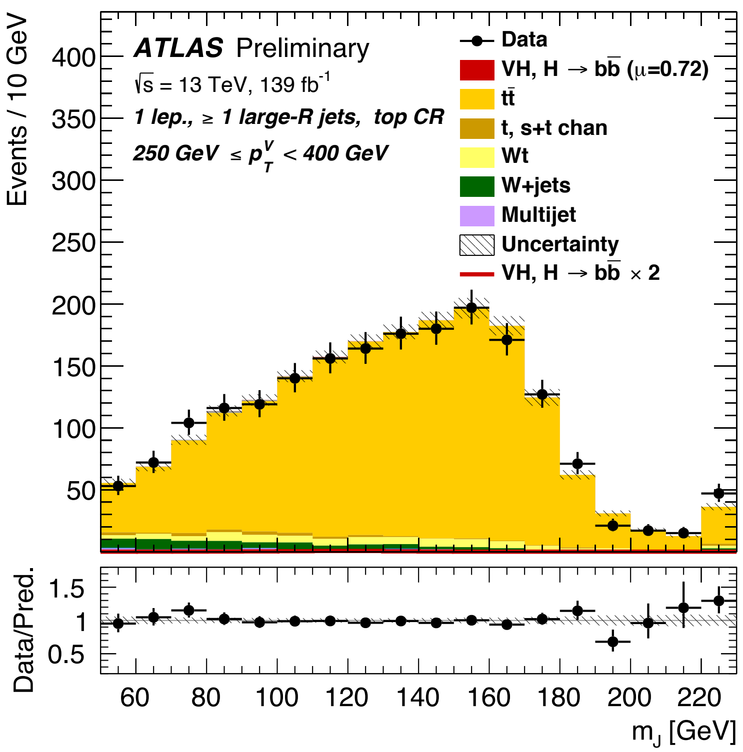
p_T^V regions

- 250-400 GeV and > 400 GeV
- Extend STXS measurement to higher p_T^V

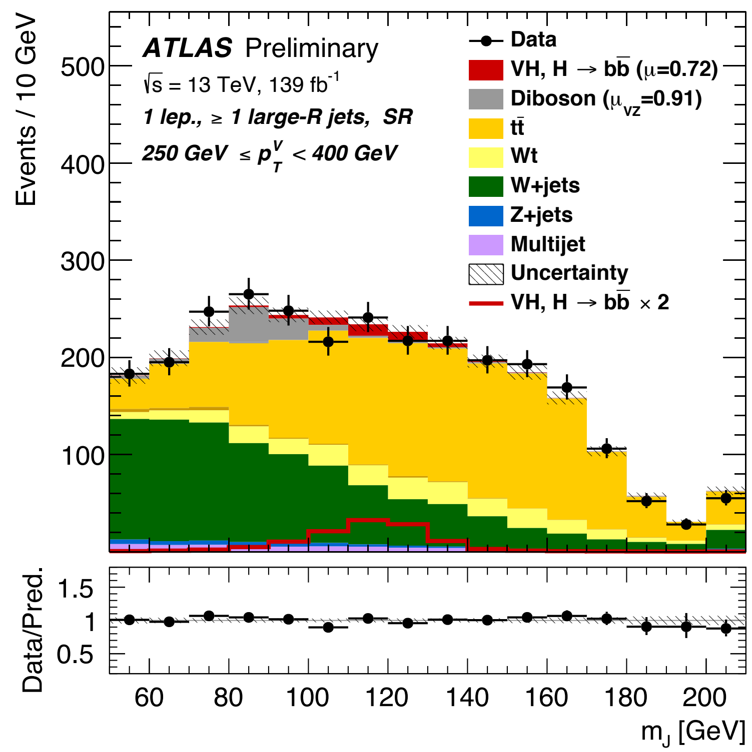
Top control region (0/1L):

- $t\bar{t}$ is the major background in 0 and 1 lepton channel
- Additional b-tagged track jets

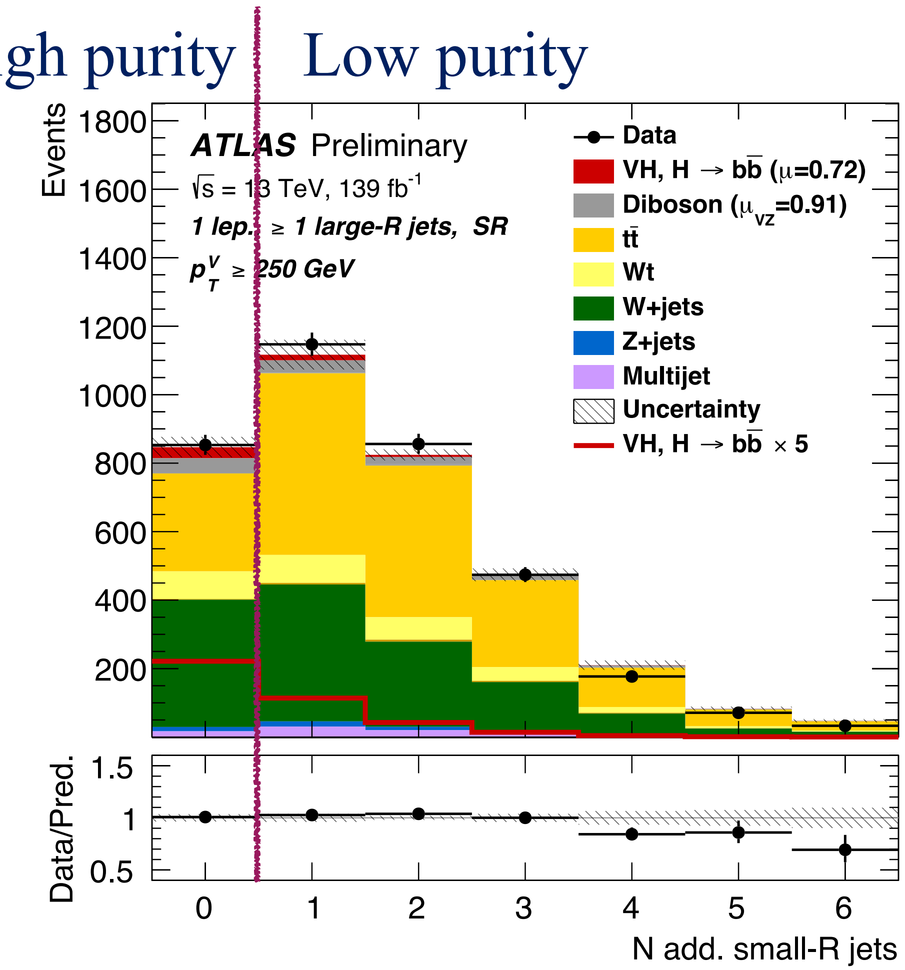
CR: ≥1



SR: 0



High purity Low purity



High/Low purity SR

- High Purity: no additional anti-k_t R=0.4 jets
- Low Purity: ≥1 additional anti-k_t R=0.4 jets

Boosted VH(bb): Results

Signal strength and significance

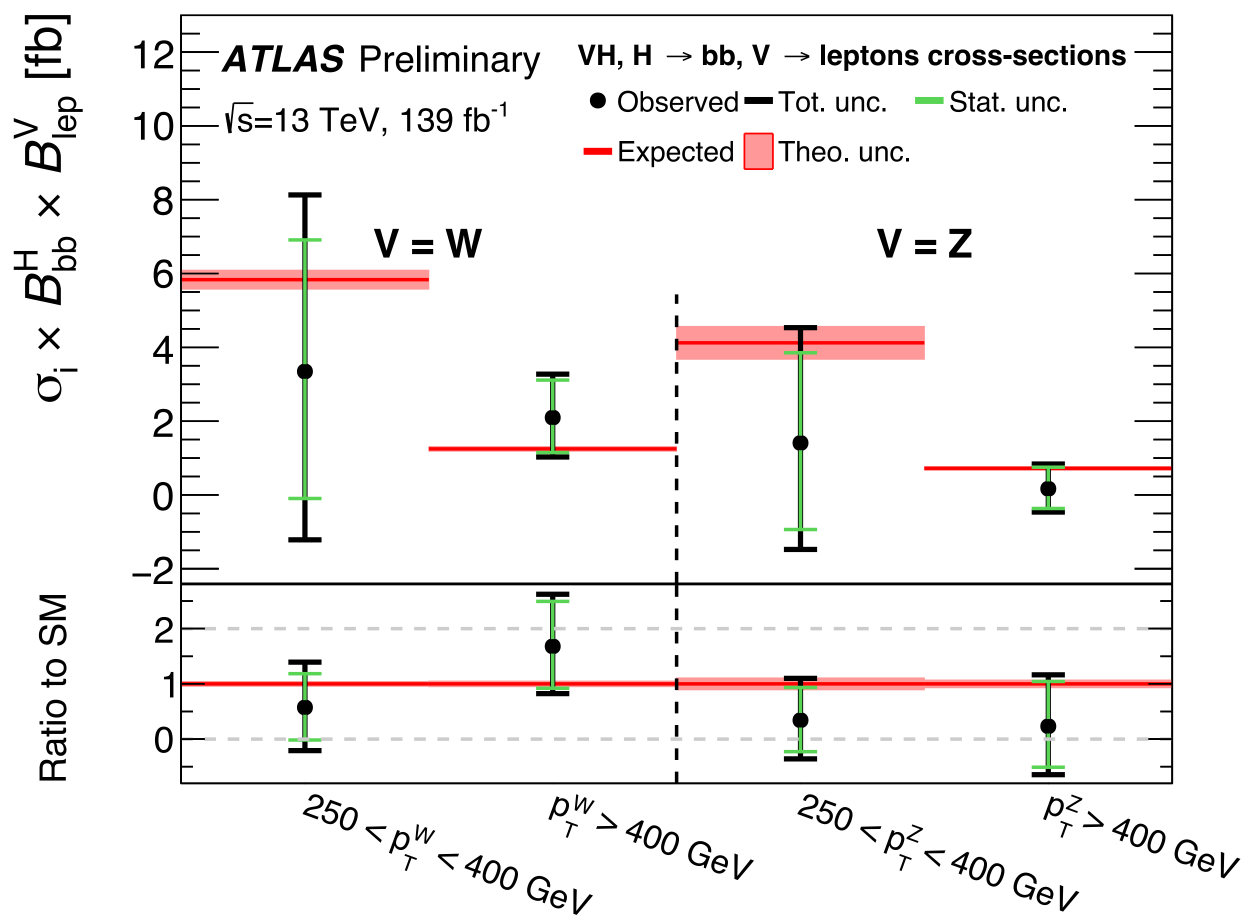
VZ Obs. (exp.) significance: 5.2 (5.7) σ

$$\mu_{VZ}^{bb} = 0.91^{+0.29}_{-0.23} = 0.91^{+0.15}_{-0.15}\,(stat.)^{+0.25}_{-0.17}\,(syst.)$$

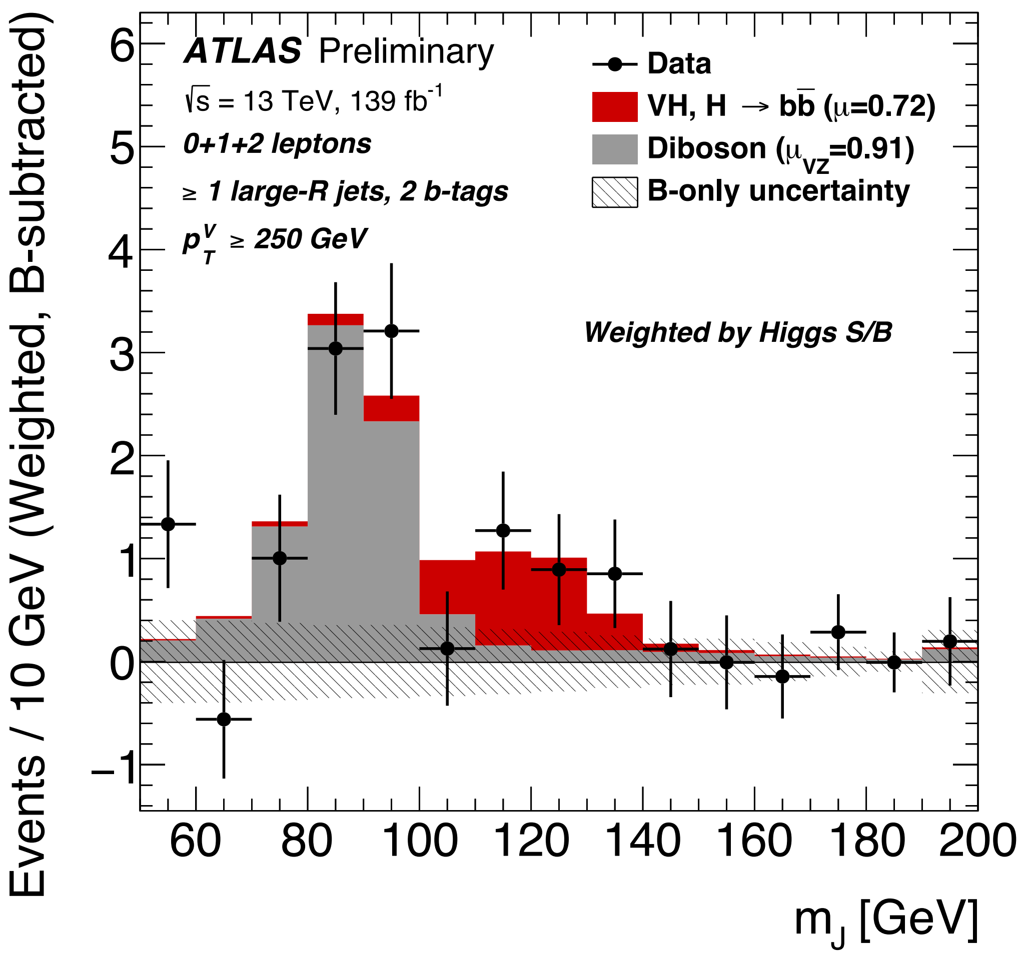
VH Obs. (exp.) significance: 2.1 (2.7) σ

$$\mu_{VH}^{bb} = 0.72^{+0.39}_{-0.36} = 0.72^{+0.29}_{-0.28}\,(stat.)^{+0.26}_{-0.22}\,(syst.)$$

Cross-section measurement



Simultaneously extracting VH(bb) and VZ(bb) signal strengths



- Good compatibility different channels / regions
- In good agreement with the SM prediction
- Analysis statistically limited

- First measurement of p_T^V> 400 GeV
- Uncertainties 70-90%

Summary

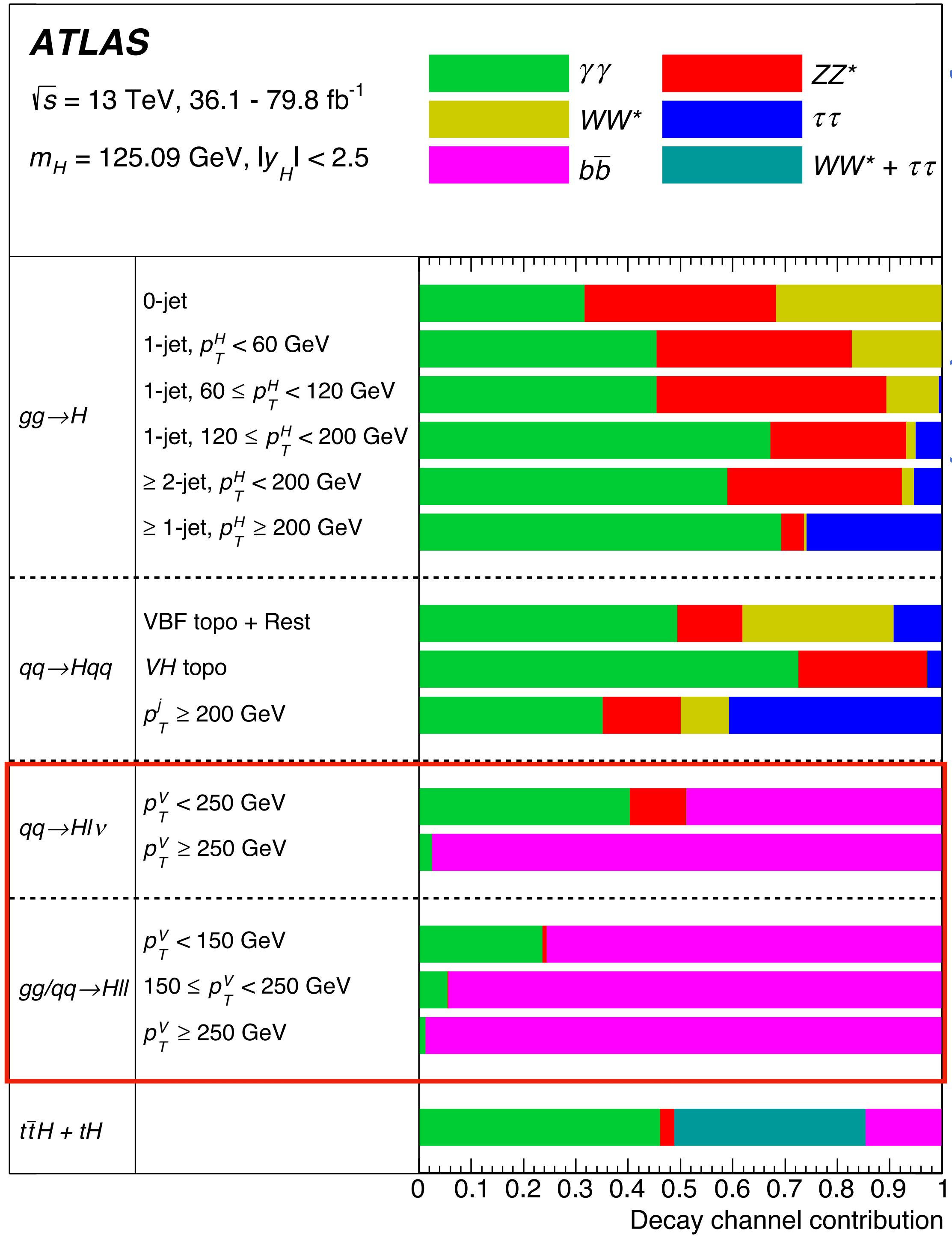
Measurement of VH(bb) process from ATLAS using the full Run-2 dataset

Two standalone analysis:

Resolved VH(bb):
- Observation of ZH(bb)
- Strong evidence of WH(bb)
- Improved STXS measurement

Boosted VH(bb):
- establish of boosted H→bb reconstruction technique
- Extend STXS bin $p_T^V > 400$ GeV

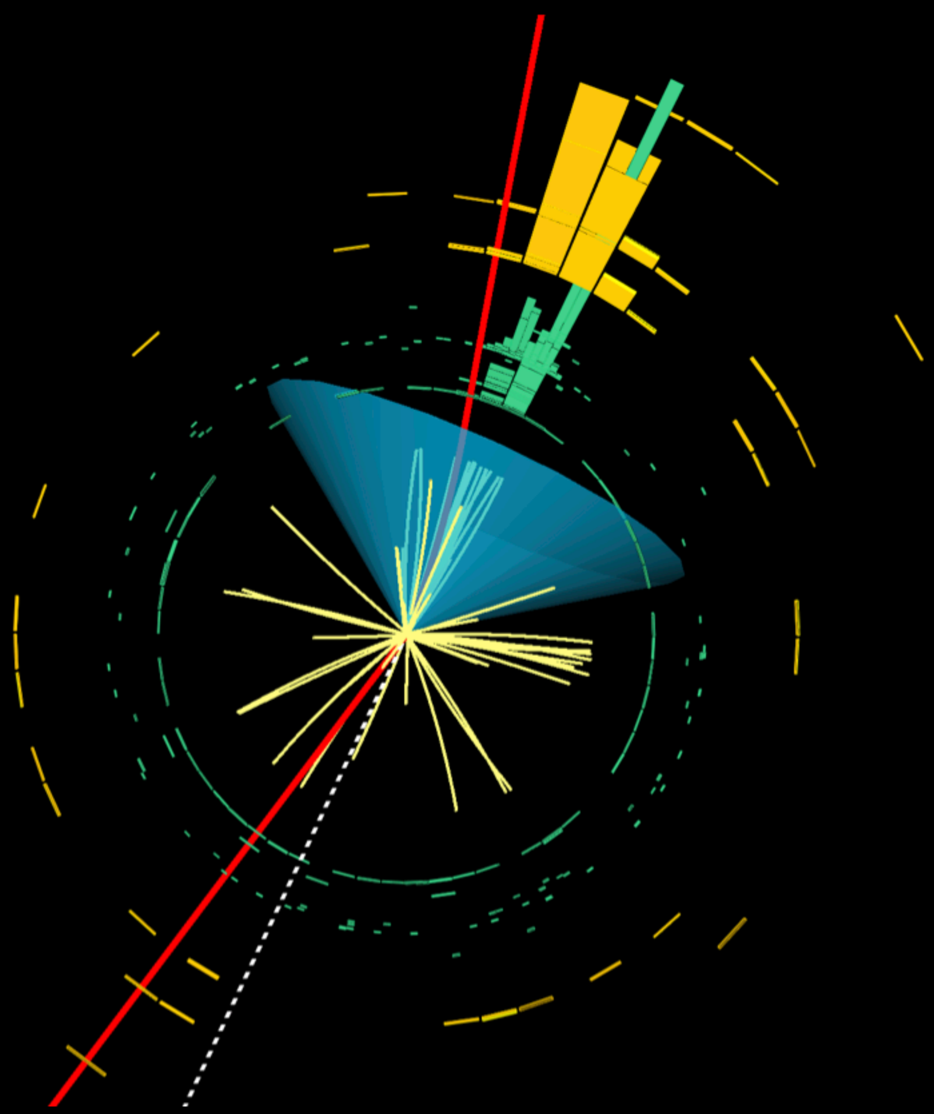
$(V)H \rightarrow b\bar{b}$ is being used as a tool to investigate Higgs production at high p_T



Phys. Rev. D 101 (2020) 012002

Contributions of the Higgs boson decay channel in the measurement of the STXS parameters

Additional material



Resolved VH(bb) Event selection

Selection	0-lepton	1-lepton: e sub-channel	1-lepton: μ sub-channel	2-lepton
Trigger	E_T^{miss}	Single lepton	E_T^{miss}	Single lepton
Leptons	0 *loose* leptons	Exactly 1 *tight* electron 0 additional *loose* leptons $p_T > 27$ GeV	Exactly 1 *tight* muon 0 additional *loose* leptons $p_T > 25$ GeV	Exactly 2 *loose* leptons $p_T > 27$ GeV Same-flavour Opposite-sign charges ($\mu\mu$)
E_T^{miss}	> 150 GeV	> 30 GeV	–	–
$m_{\ell\ell}$	–	–	–	81 GeV $< m_{\ell\ell} <$ 101 GeV
Jet p_T	> 20 GeV for $\lvert\eta\rvert < 2.5$ > 30 GeV for $2.5 < \lvert\eta\rvert < 4.5$			
b-jets	Exactly 2 b-tagged jets			
Leading b-tagged jet p_T	> 45 GeV			
Jet categories	Exactly 2 / Exactly 3 jets	Exactly 2 / Exactly 3 jets		Exactly 2 / ≥ 3 jets
H_T	> 120 GeV (2 jets), >150 GeV (3 jets)	–		–
$\min[\Delta\phi(\vec{E}_T^{miss}, \vec{jets})]$	$> 20^\circ$ (2 jets), $> 30^\circ$ (3 jets)	–		–
$\Delta\phi(\vec{E}_T^{miss}, \vec{bb})$	$> 120^\circ$	–		–
$\Delta\phi(\vec{b_1}, \vec{b_2})$	$< 140^\circ$	–		–
$\Delta\phi(\vec{E}_T^{miss}, \vec{p}_T^{miss})$	$< 90^\circ$	–		–
p_T^V regions	– 150 GeV $< p_T^V <$ 250 GeV $p_T^V > 250$ GeV	– 150 GeV $< p_T^V <$ 250 GeV $p_T^V > 250$ GeV		75 GeV $< p_T^V <$ 150 GeV 150 GeV $< p_T^V <$ 250 GeV $p_T^V > 250$ GeV
Signal regions	$\Delta R(\vec{b_1}, \vec{b_2})$ signal selection			
Control regions	High and low $\Delta R(\vec{b_1}, \vec{b_2})$ side-bands			

Boosted VH(bb) 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
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 250GeV	> 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, $\lvert\eta\rvert < 2.5$, associated to the leading large-R jet				
b-jets	leading two track-jets associated to the leading large-R must be b-tagged (MV2c10, 70%)				
m_{J}	> 50 GeV				
$\min[\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets$)]$	$> 30^\circ$	-			
$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, H_{\mathrm{cand}})$	$> 120^\circ$	-			
$\Delta\phi\ (\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{E}_{\mathrm{T,\ trk}}^{\mathrm{miss}})$	$< 90^\circ$	-			
$\Delta y(V, H_{\mathrm{cand}})$	-	$\lvert\Delta y(V, H_{\mathrm{cand}})\rvert < 1.4$			
$m_{\ell\ell}$	-			66 GeV $< m_{\ell\ell} <$ 116 GeV	
Lepton p_{T} imbalance	-			$(p_{\mathrm{T}}^{\ell_1} - p_{\mathrm{T}}^{\ell_2})/p_{\mathrm{T}}^{Z} < 0.8$	

Boosted VH(bb): Large-R jet mass resolution

Final discriminant: Mass of the large-R jet

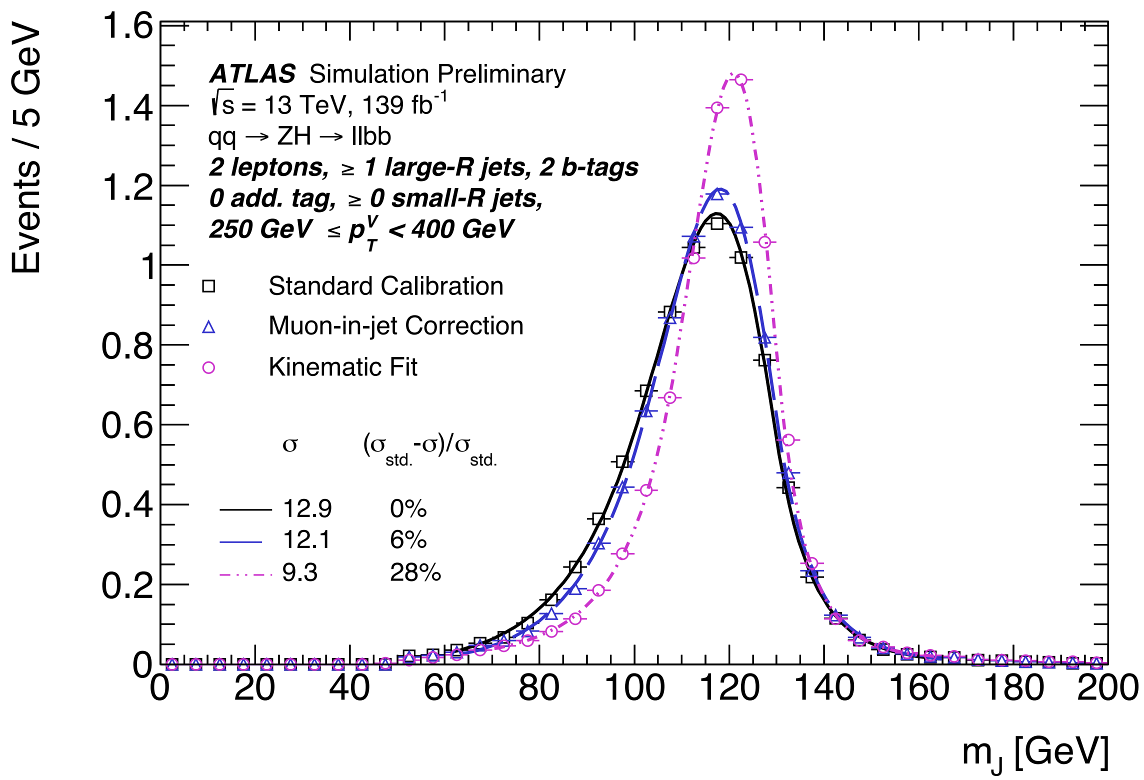
0/1L : muon-in-jet correction

- Correct for semi-muonic heavy hadron decays
- ~10% (~4%) improvement on mass resolution (significance)

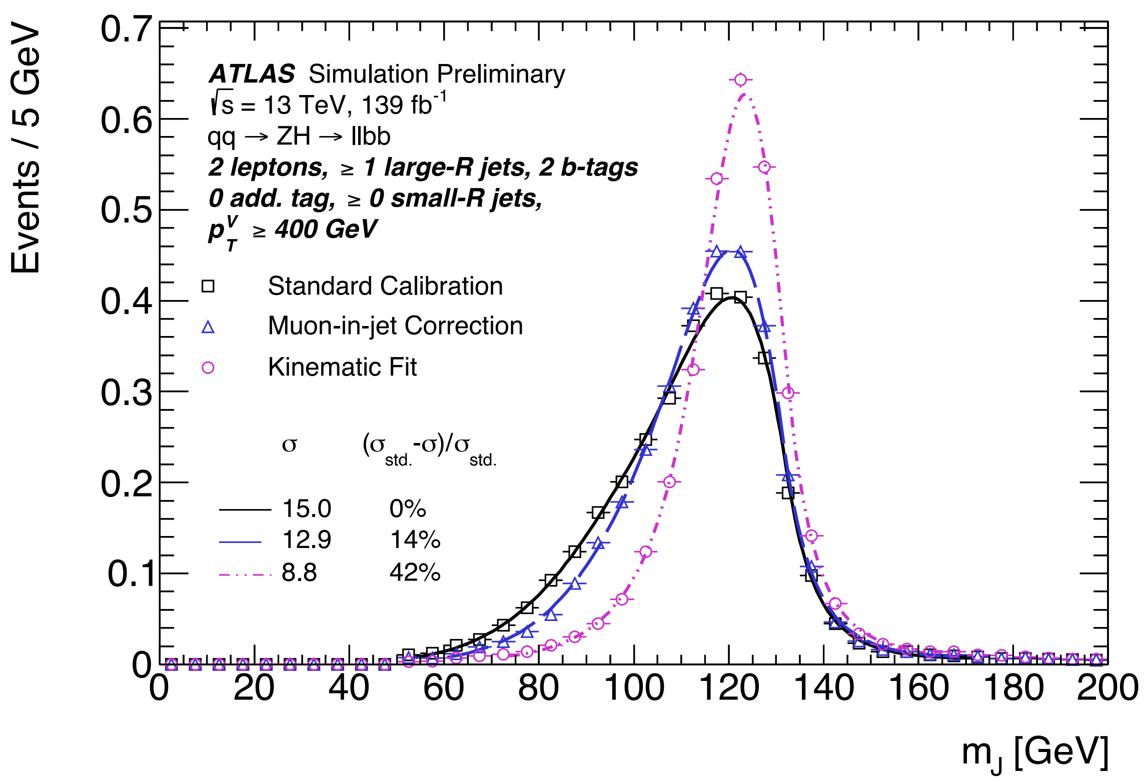
2L : kinematic fit

- exploit excellent energy resolution of leptons
- ~30% (~14%) improvement on mass resolution (significance)

Similar resolution as m_{bb} in resolved analysis



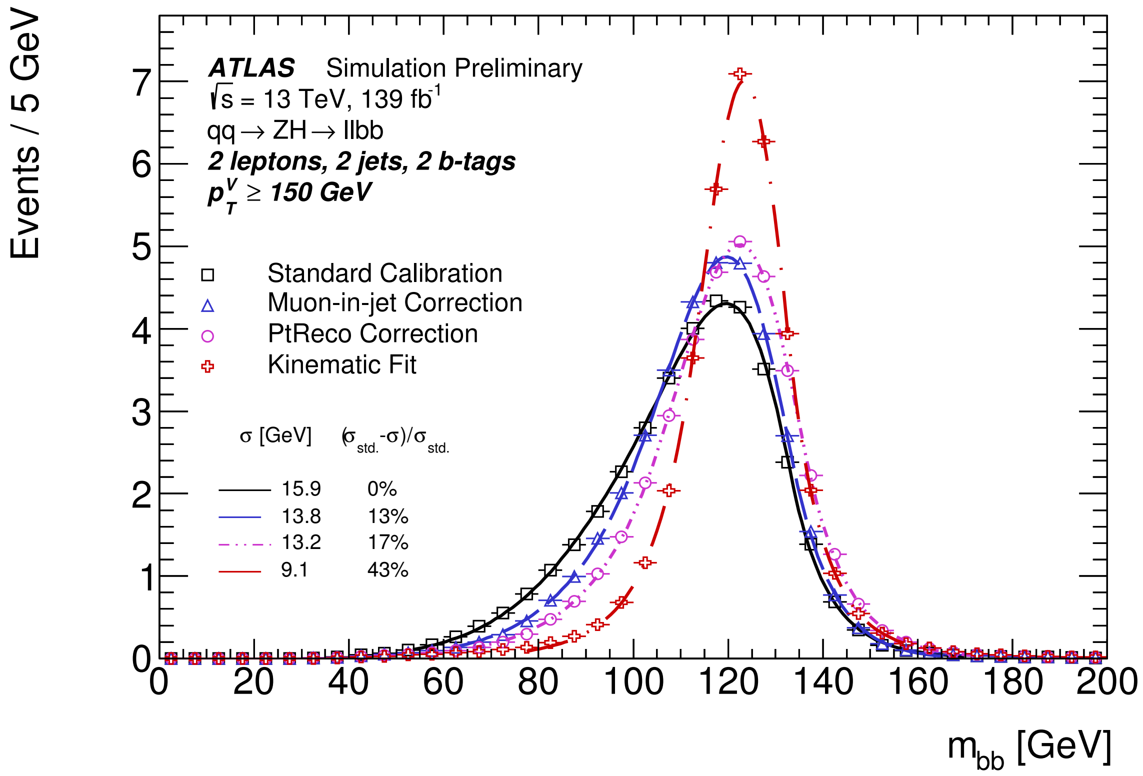
$250 \text{ GeV} < p_T^Z < 400 \text{ GeV}$



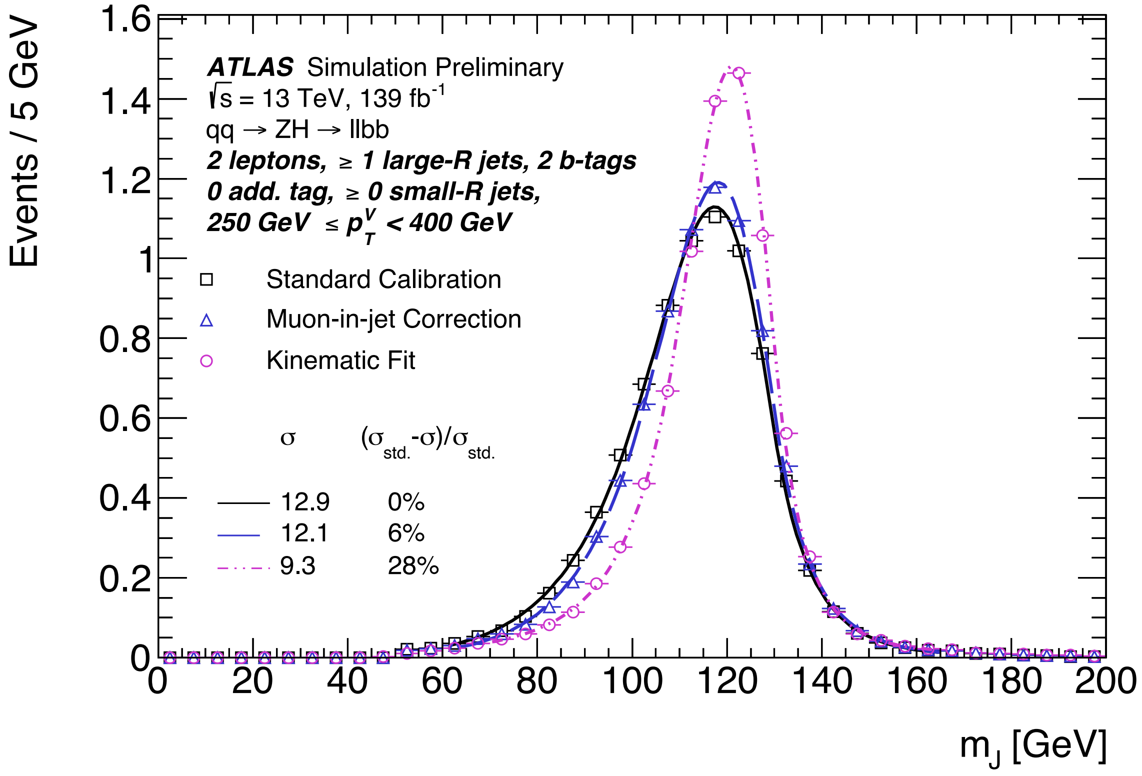
$p_T^Z > 400 \text{ GeV}$

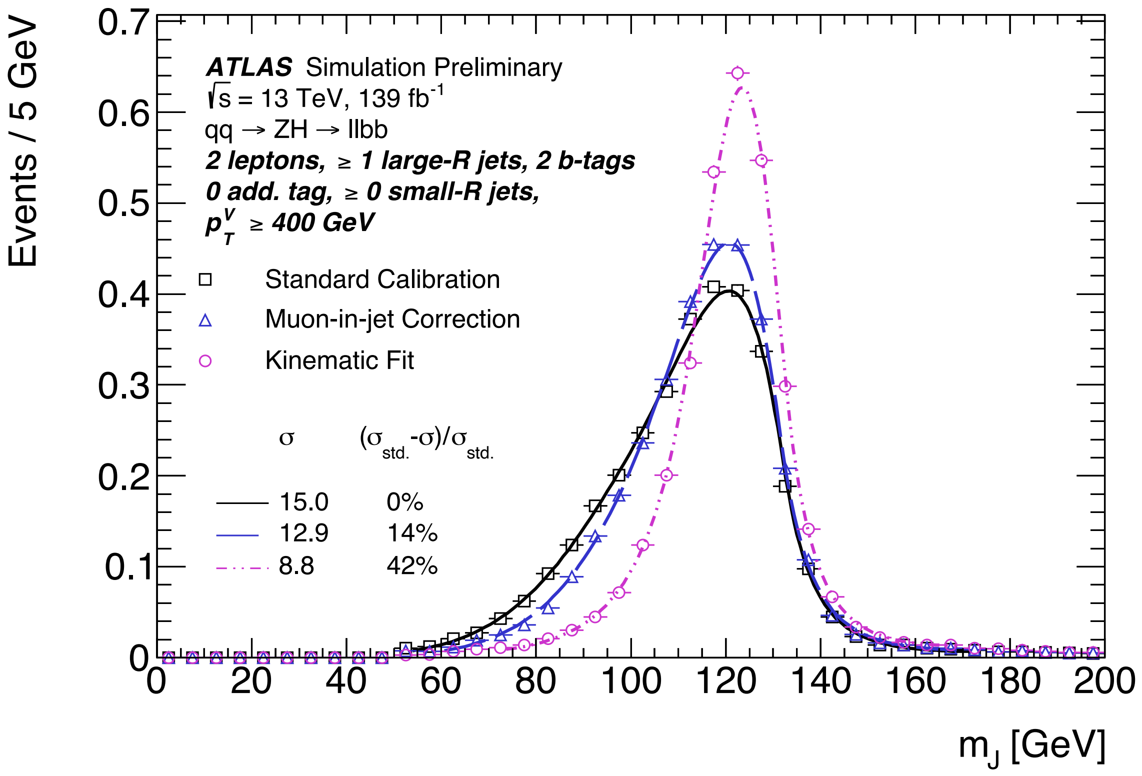
B-jet energy correction

Resolved mbb

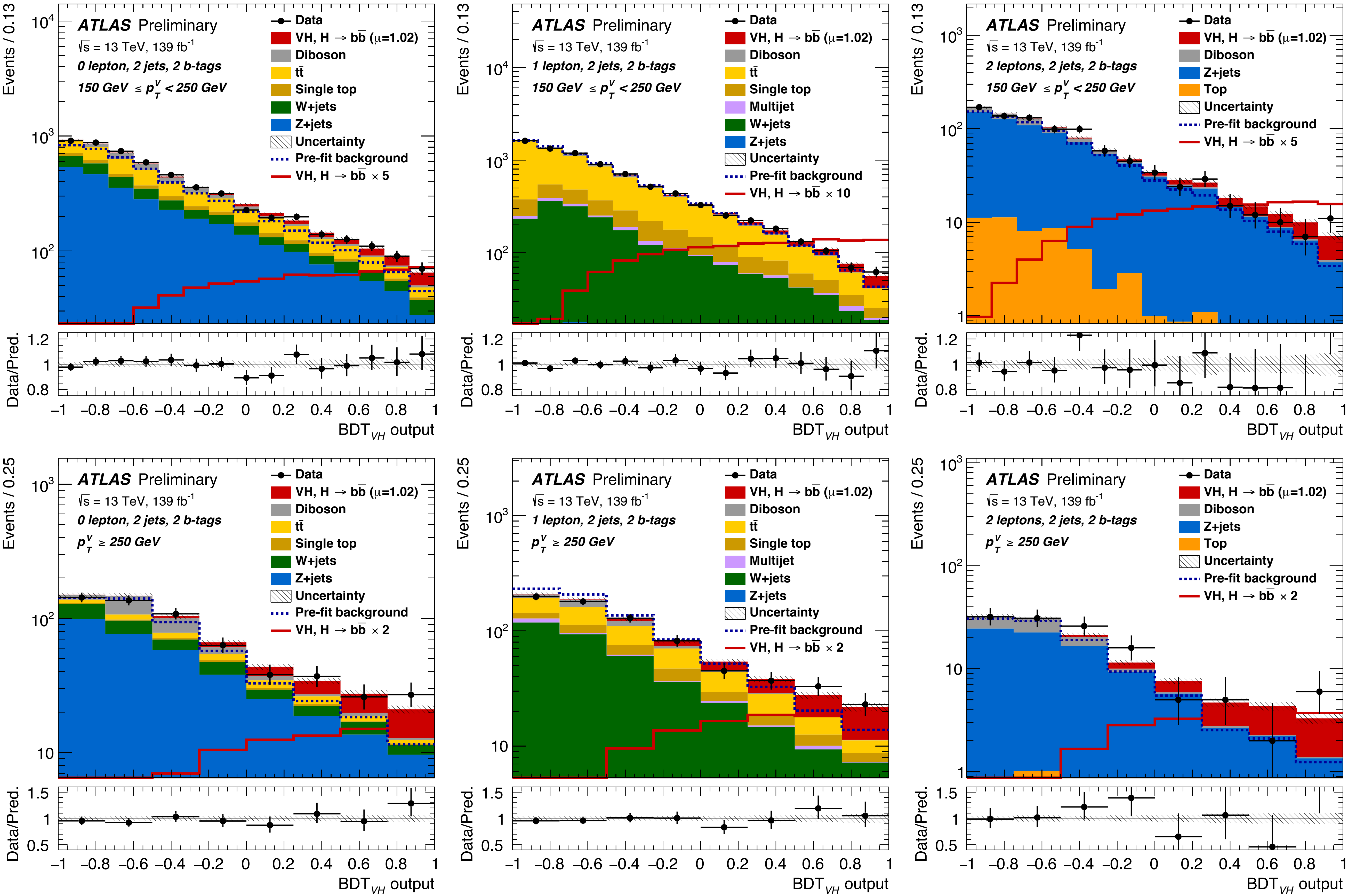


Boosted mass of large-R jet

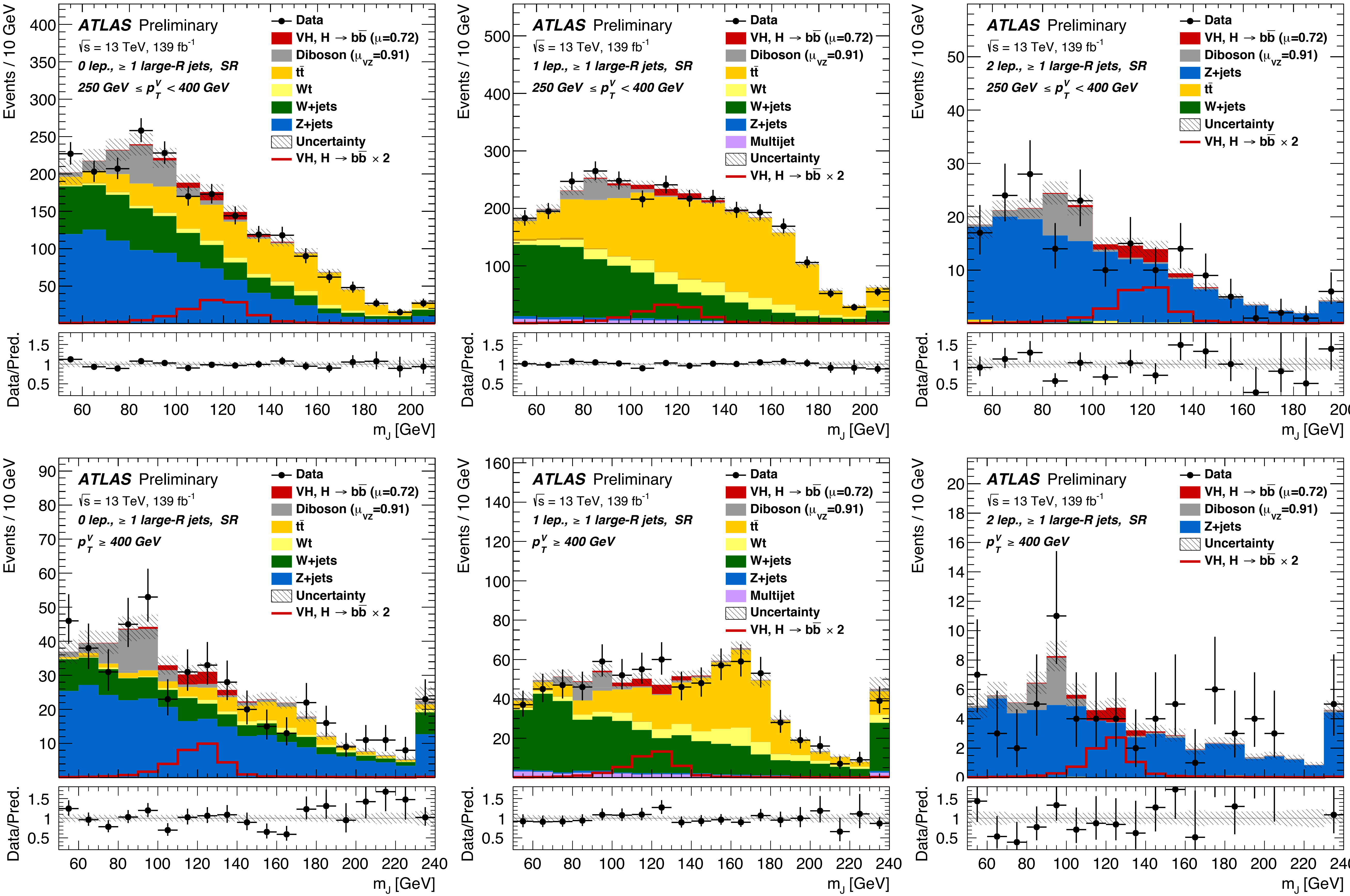




Resolved VH(bb): BDT output plots



Boosted VH(bb): Mass of large-R jet



Resolved VH(bb): The fit model

- Perform a binned likelihood fit simultaneously in the 14 signal regions and 28 control regions

14 SR / 28 CR

Channel	Region	75 GeV~150 GeV 2 jet	75 GeV~150 GeV 3+jets	150 GeV ~ 250 GeV 2 jet	150 GeV ~ 250 GeV 3(+)jets	> 250 GeV 2 jet	> 250 GeV 3(+)jets
	Low ΔR CR			Yields	Yields	Yields	Yields
0-lepton	SR			BDT	BDT	BDT	BDT
	High ΔR CR			Yields	Yields	Yields	Yields
	Low ΔR CR			Yields	Yields	Yields	Yields
1-lepton	SR			BDT	BDT	BDT	BDT
	High ΔR CR			Yields	Yields	Yields	Yields
	Low ΔR CR			Yields	Yields	Yields	Yields
2-lepton	SR	BDT	BDT	BDT	BDT	BDT	BDT
	High ΔR CR			Yields	Yields	Yields	Yields

Normalisation factors:

Measurement region ($\|y_H\| < 2.5$, $H \to b\bar{b}$)	SM prediction [fb]	Result [fb]	Stat. unc. [fb]	Syst. unc. [fb] Th. sig.	Syst. unc. [fb] Th. bkg.	Syst. unc. [fb] Exp.
$W \to \ell\nu$; $150 < p_T^{W,t} < 250$ GeV	24.0 ± 1.1	19.0 ± 12.1	± 7.7	± 0.9	± 5.5	± 6.0
$W \to \ell\nu$; $p_T^{W,t} > 250$ GeV	7.1 ± 0.3	7.2 ± 2.2	± 1.9	± 0.4	± 0.8	± 0.7
$Z \to \ell\ell, \nu\nu$; $75 < p_T^{Z,t} < 150$ GeV	50.6 ± 4.1	42.5 ± 35.9	± 25.3	± 5.6	± 17.2	± 19.7
$Z \to \ell\ell, \nu\nu$; $150 < p_T^{Z,t} < 250$ GeV	18.8 ± 2.4	20.5 ± 6.2	± 5.0	± 2.3	± 2.4	± 2.3
$Z \to \ell\ell, \nu\nu$; $p_T^{Z,t} > 250$ GeV	4.9 ± 0.5	5.4 ± 1.7	± 1.5	± 0.5	± 0.5	± 0.3

Boosted VH(bb): The fit model

Background modelling

Same strategy as resolved analysis:

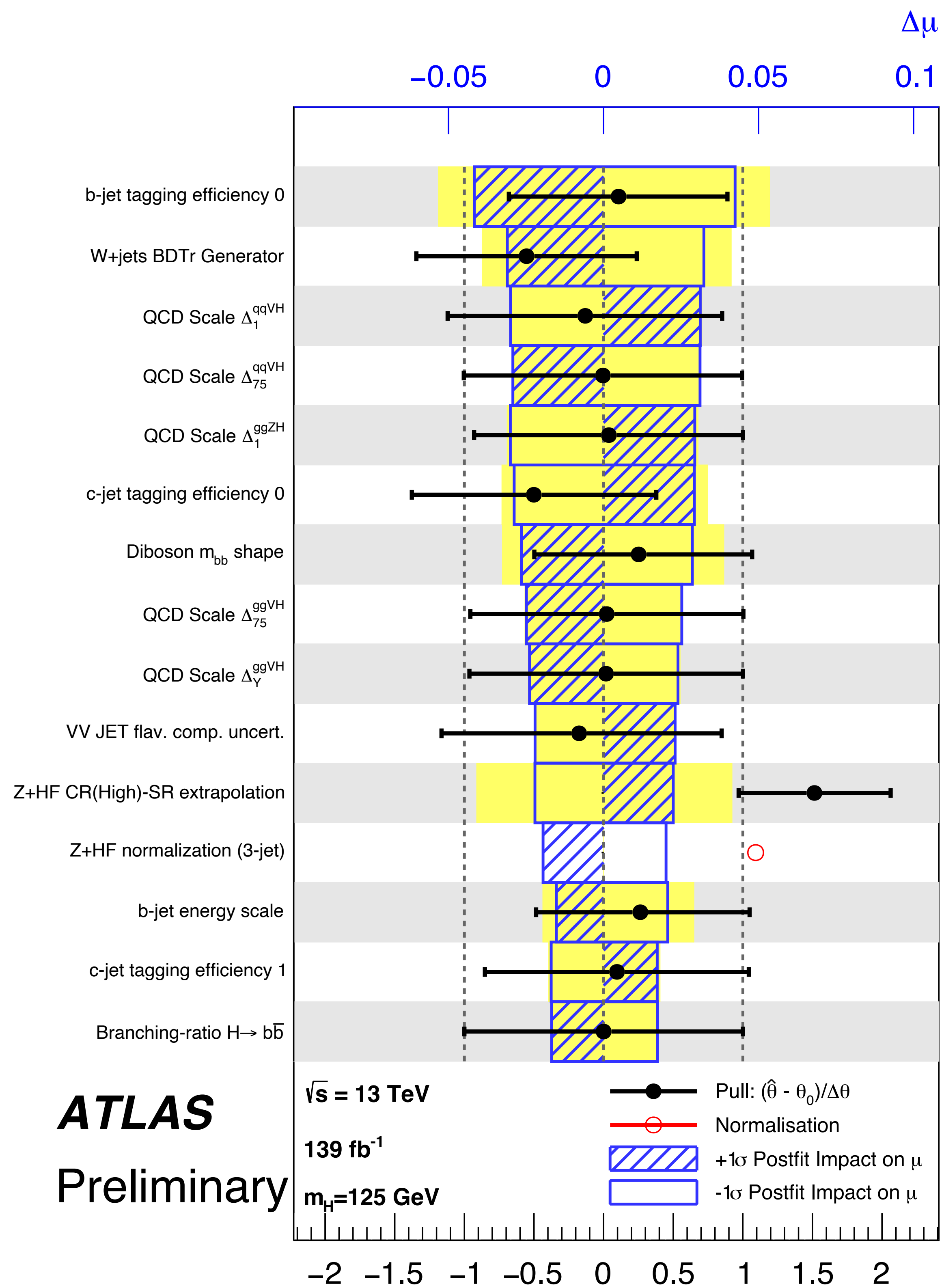
- Normalisations of the main backgrounds floated within the fit ($t\bar{t}$, W+jets, Z+jets)
- Acceptance and shape uncertainties are derived from comparisons between simulated samples

The fit model

- Binned profile likelihood fit in the mass of large-R jet distributions in 14 regions
- **Simultaneously extracting VH(bb) and VZ(bb) signal strengths**

Channel	Categories					
	$250 < p_{\mathrm{T}}^{V} < 400$ GeV			$p_{\mathrm{T}}^{V} \geq 400$ GeV		
	0 add. b-track-jets		≥ 1 add. b-track-jets	0 add. b-track-jets		≥ 1 add. b-track-jets
	0 add. small-R jets	≥ 1 add. small-R jets		0 add. small-R jets	≥ 1 add. small-R jets	
0-lepton	HP SR	LP SR	CR	HP SR	LP SR	CR
1-lepton	HP SR	LP SR	CR	HP SR	LP SR	CR
2-lepton	SR			SR		

Resolved VH(bb): Systematics ranking and breakdown



Source of uncertainty	σ_μ VH	WH	ZH
Total	0.177	0.260	0.240
Statistical	0.115	0.182	0.171
Systematic	0.134	0.186	0.168
Statistical uncertainties			
Data statistical	0.108	0.171	0.157
$t\bar{t}$ $e\mu$ control region	0.014	0.003	0.026
Floating normalisations	0.034	0.061	0.045
Experimental uncertainties			
Jets	0.043	0.050	0.057
E_T^{miss}	0.015	0.045	0.013
Leptons	0.004	0.015	0.005
b-tagging: b-jets	0.045	0.025	0.064
b-tagging: c-jets	0.035	0.068	0.010
b-tagging: light-flavour jets	0.009	0.004	0.014
Pile-up	0.003	0.002	0.007
Luminosity	0.016	0.016	0.016
Theoretical and modelling uncertainties			
Signal	0.052	0.048	0.072
Z + jets	0.032	0.013	0.059
W + jets	0.040	0.079	0.009
$t\bar{t}$	0.021	0.046	0.029
Single top quark	0.019	0.048	0.015
Diboson	0.033	0.033	0.039
Multi-jet	0.005	0.017	0.005
MC statistical	0.031	0.055	0.038

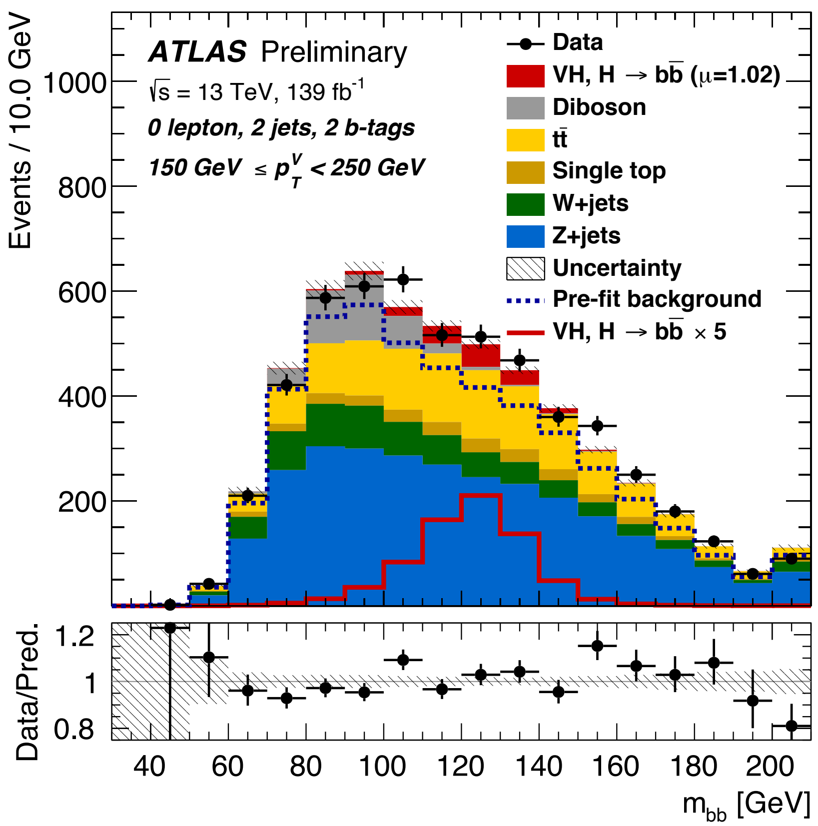
Boosted VH(bb): Systematics breakdown

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_T^{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
$\hookrightarrow$ Z + jets		0.048
$\hookrightarrow$ W + jets		0.058
$\hookrightarrow$ $t\bar{t}$		0.035
$\hookrightarrow$ Single top quark		0.027
$\hookrightarrow$ Diboson		0.032
$\hookrightarrow$ Multijet		0.009
MC statistical		0.092

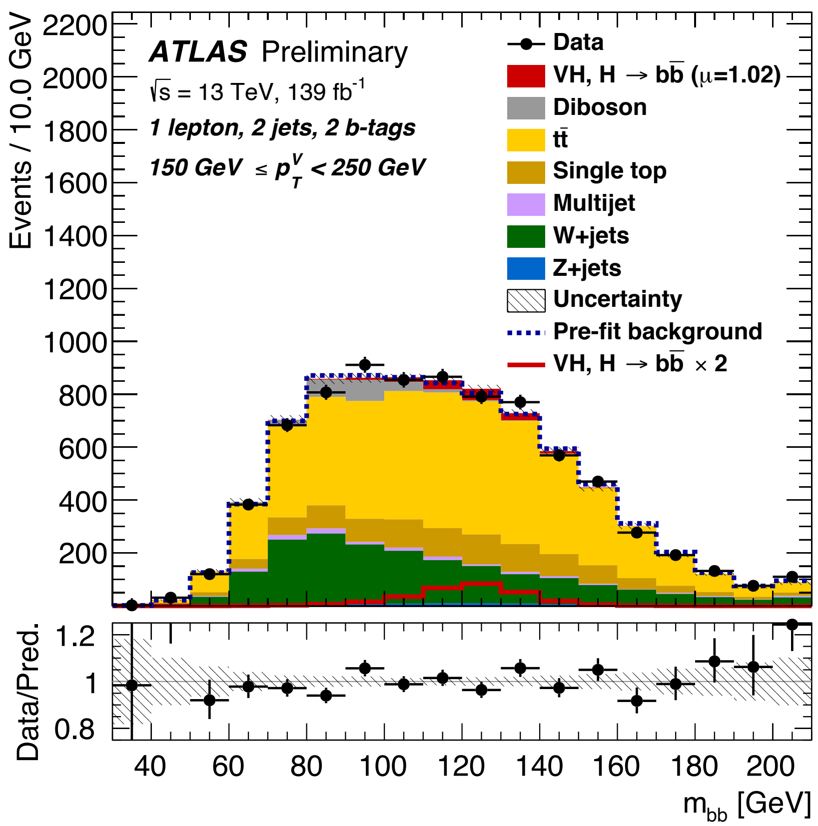
Resolved VH(bb): Background modelling

Dominant backgrounds

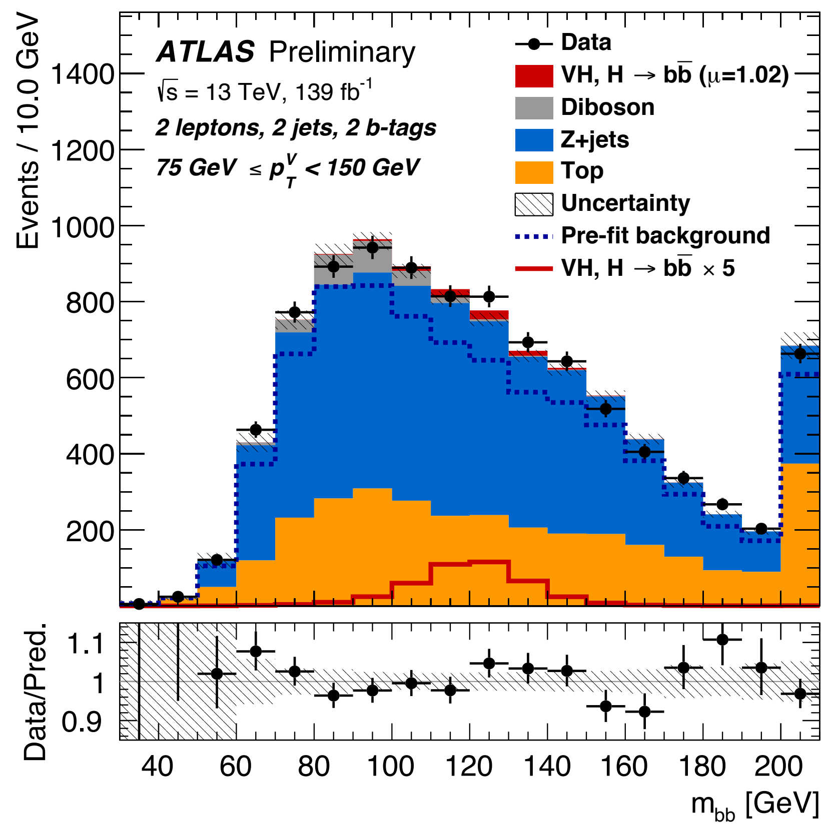
0L: W/Z+jets, $t\bar{t}$



1L: W+jets, $t\bar{t}$



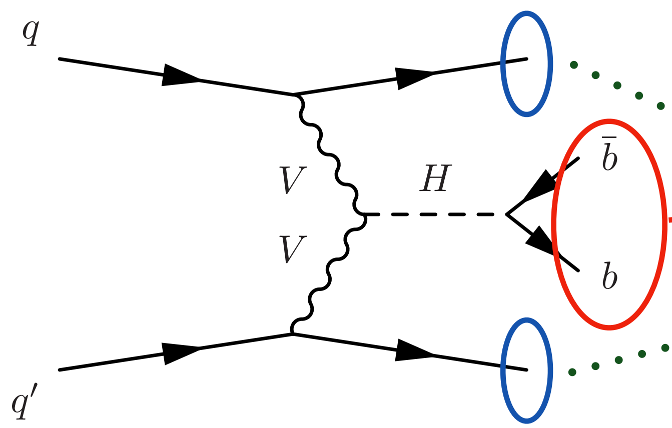
2L: Z+jets, Top

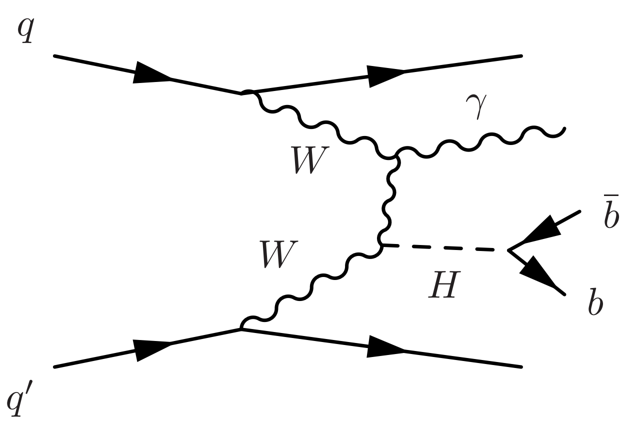


Background estimation

- Data-driven method for 2L top background: use $e\mu$ control region data
- Normalisations of dominant backgrounds are **floated** in the final fit to data
- Acceptance and shape uncertainties:
  - Z+jets shape uncertainties on m_{bb} and p_T^V are estimated from an m_{bb}-sideband region in data
  - Other uncertainties are derived from comparisons between simulated samples

VBF and qqH(→ bb)γ





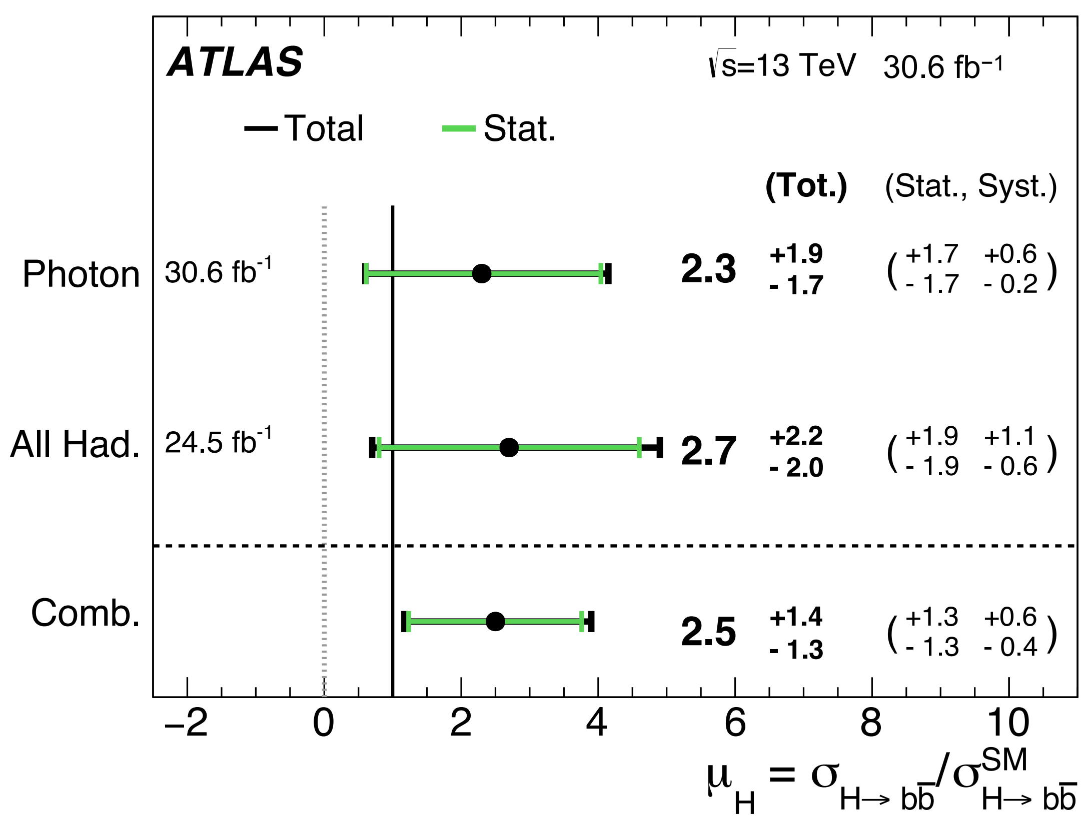
Event Selection

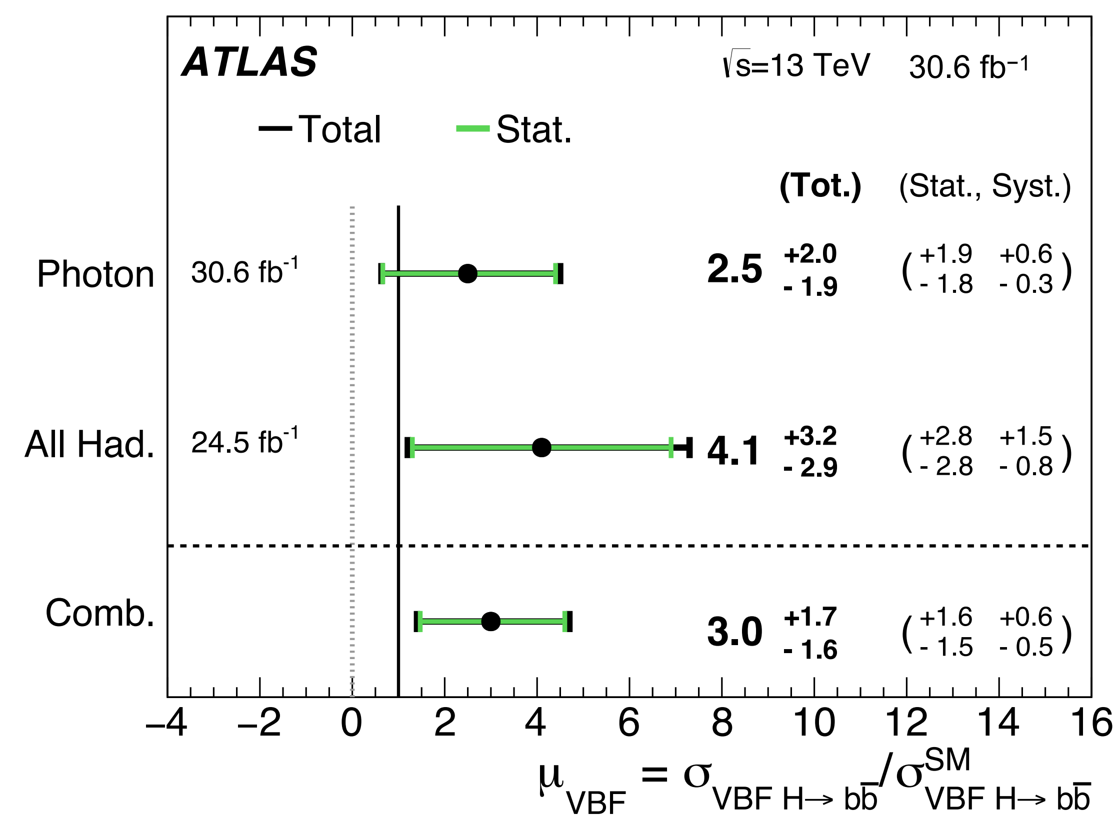
- Higgs candidate: 2 b-tagged anti-k_t R=0.4 jets
- VBF jets: two light-quark jets with a large rapidity gap

Three channels

- Two-central channel
- Four-central channel
- Photon channel

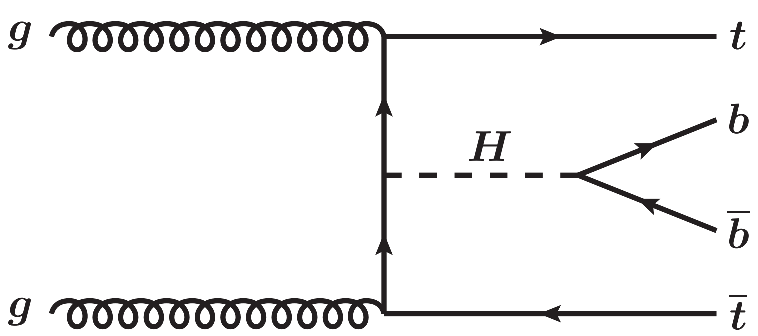
- MVA(BDT) is used to define event categories of varying signal purity
- Perform a binned likelihood fit on the m_{bb} spectrum

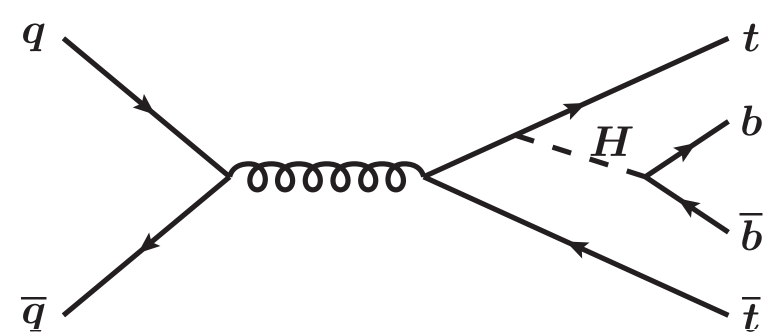




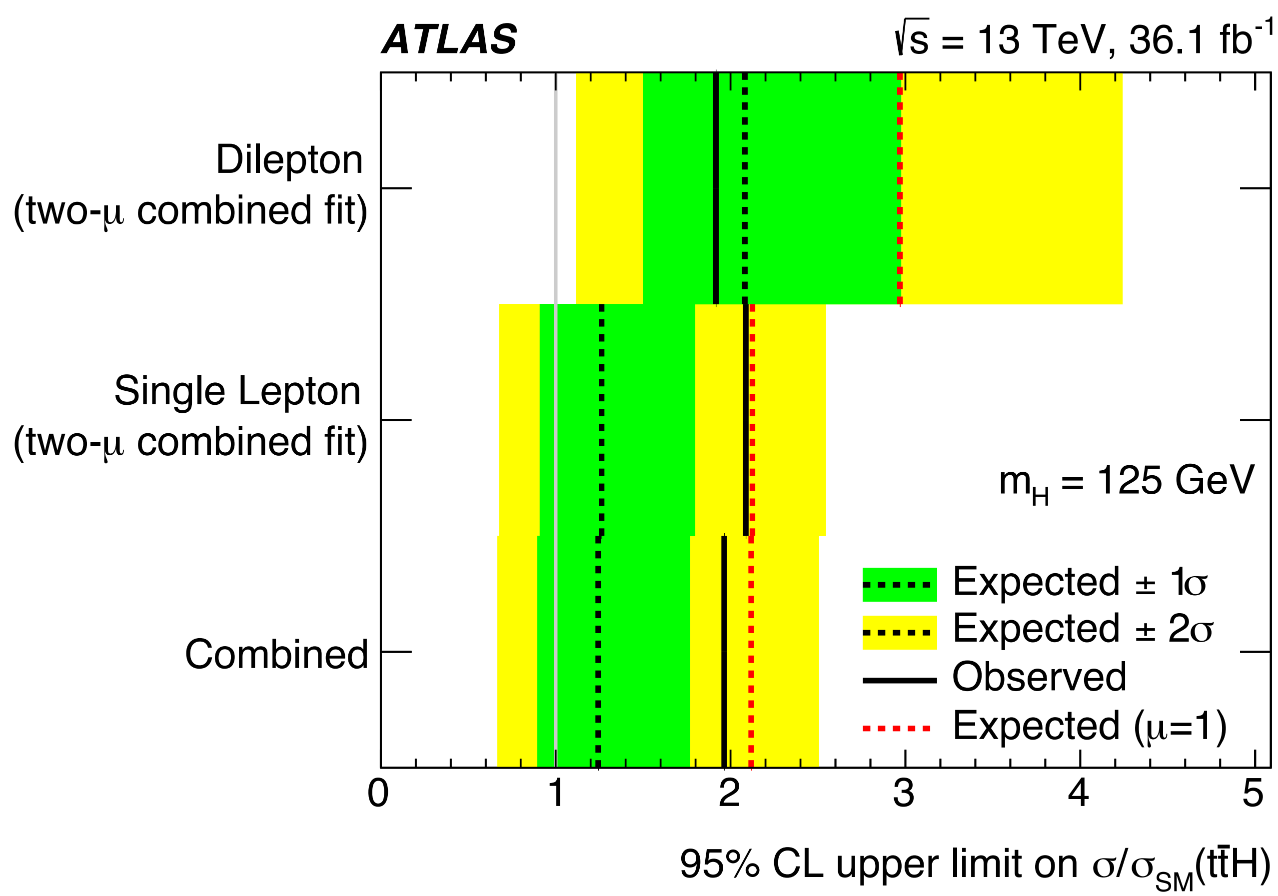
Phys. Rev. D 98 (2018) 052003

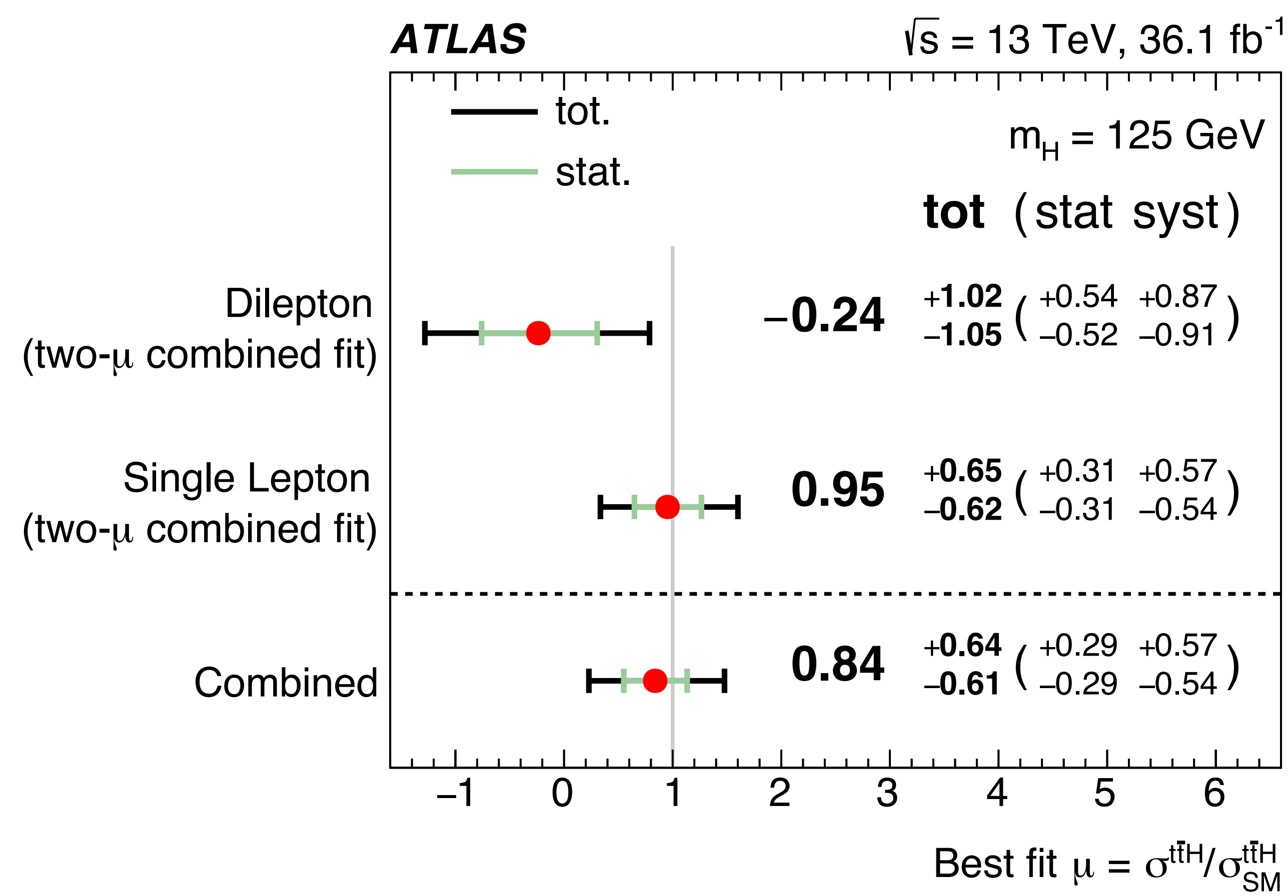
ttH+tH, H(bb)





- Single-lepton and dilepton channel a
- Events are categorised using number of jets and b-tagging discriminant
- Using BDT output as the final discriminant





Phys. Rev. D 97 (2018) 072016