Searches for Higgs coupling to charm quarks at the ATLAS experiment

Greta Brianti on behalf of the ATLAS collaboration

Higgs Hunting 2025

15.07.2025

Introduction

Charm Yukawa coupling:

Many theories predict modifications to the earlier-generation quark Yukawa couplings. Of these, the charm coupling is the most accessible one.

Experimental challenges:

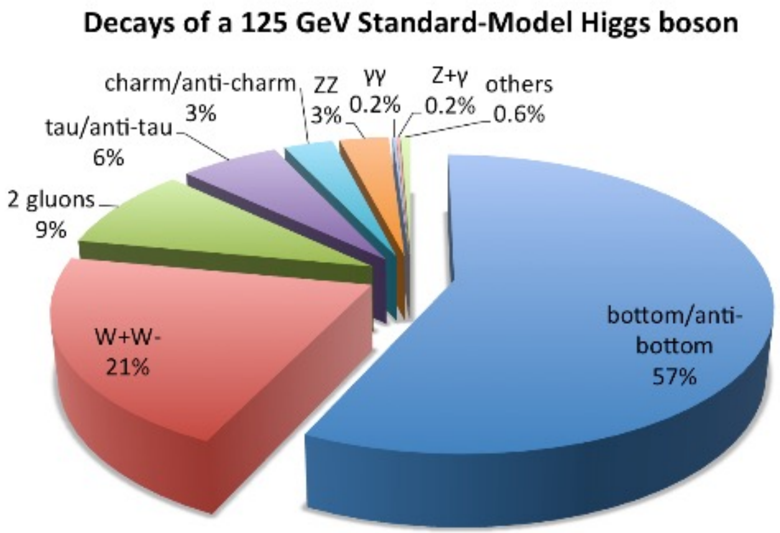
- **Standard Model Branching Ratio ~ 3%**
- **c-jet identification**
- High background rate

Direct search: ATLAS VH(bb/cc) Run 2 Analysis

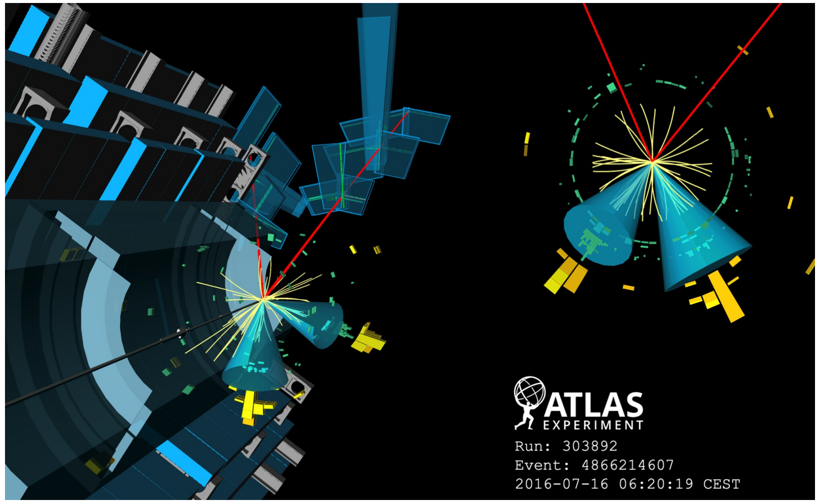
- Different possible channels through selecting the vector boson decay
- QCD background suppression

Indirect search: ATLAS H(γγ) + c production

- Exploiting production cross-section dependence from Yukawa coupling to c-quarks



$Z(\mu\mu)H(cc)$

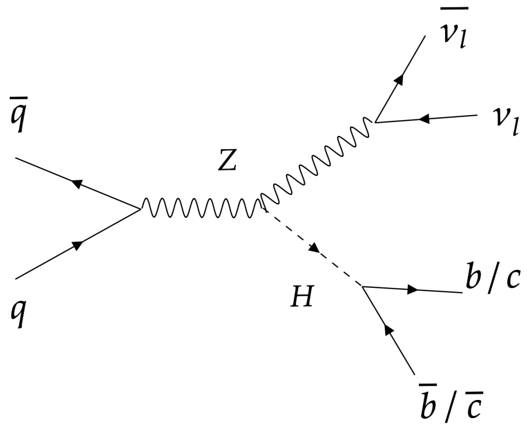


VH(bb/cc) Run 2 Analysis

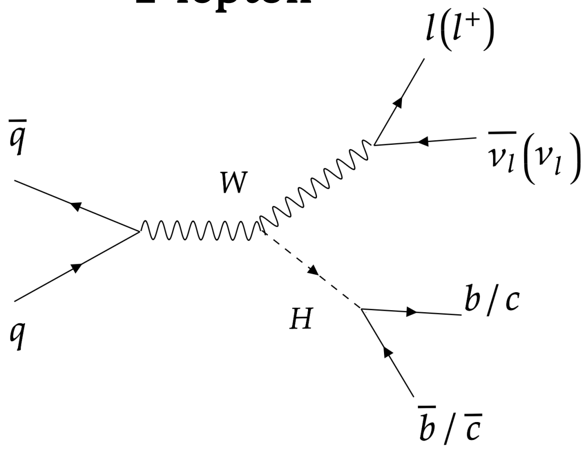
➢ Three channels

- $Z(\nu\nu)\ +\ H(bb/cc)$
- $W(l\nu_l)\ +\ H(bb/cc)$
- $Z(l^+l^-)\ +\ H(bb/cc)$

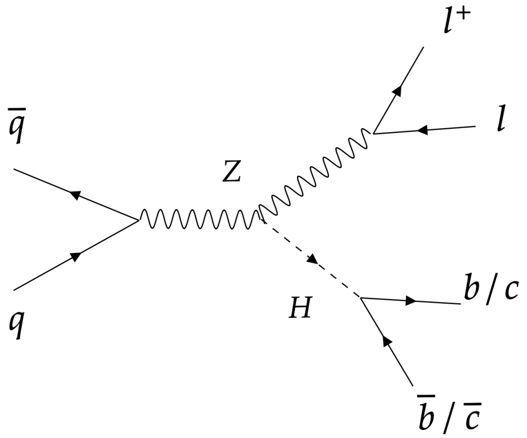
0-lepton



1-lepton

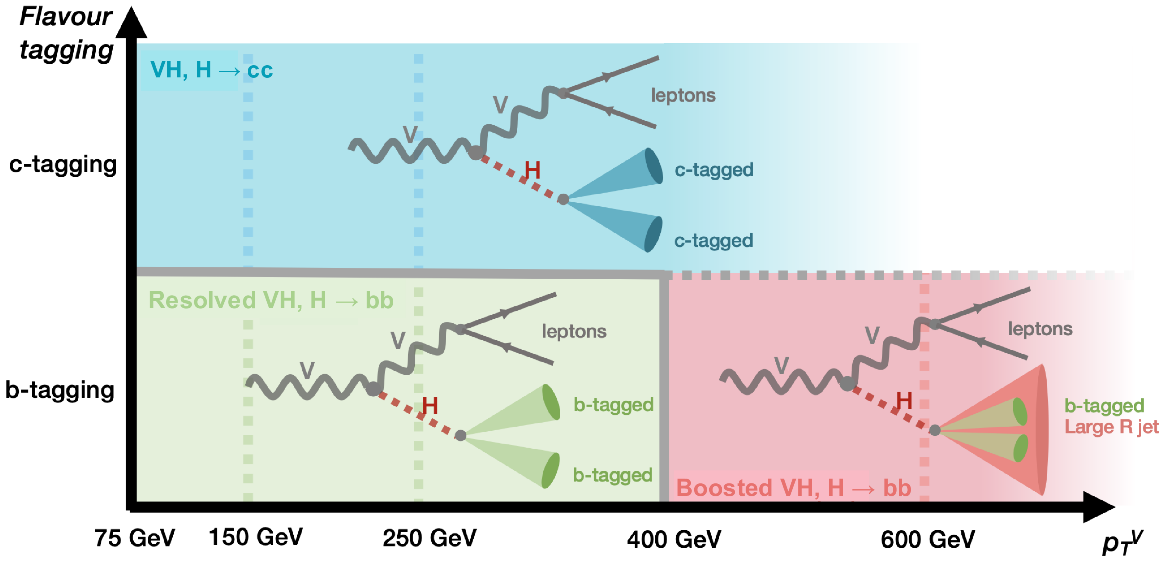


2-lepton



➢ Topology

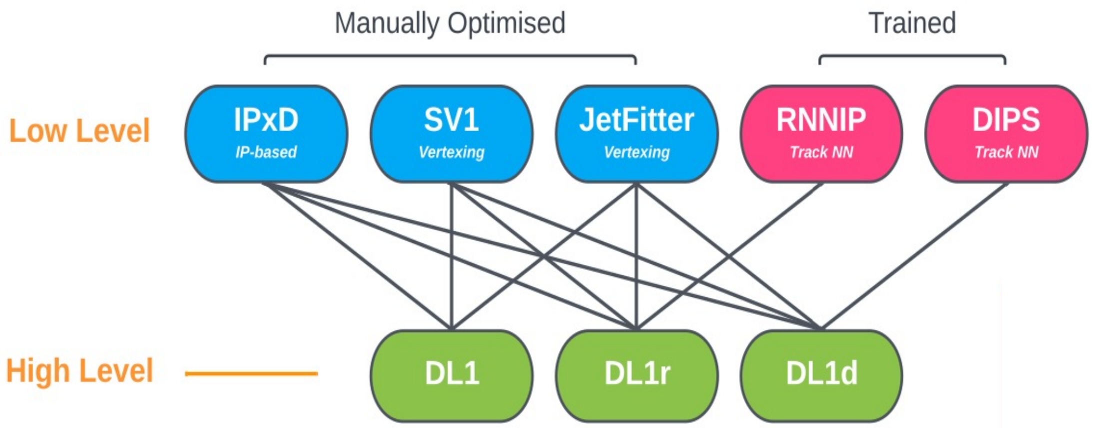
- **Exploit similarities between VH(bb) and VH(cc) topologies for the full Run2 dataset**
- <u>Resolved</u> and <u>boosted</u> topology considered for VH(bb)
- Orthogonality ensured with $p_T^V > 400$ GeV
- Resolved reconstruction only for VH(cc)



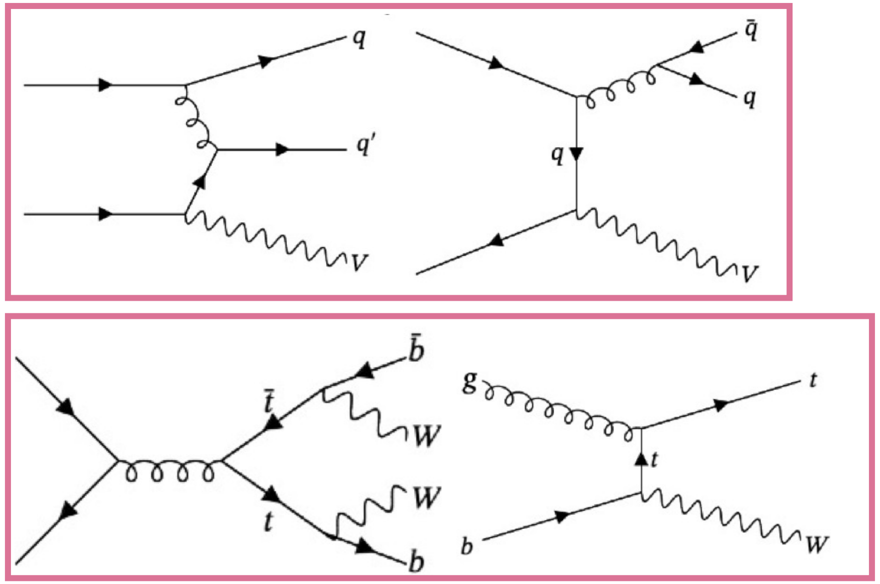
Resolved topology

[S. Van Stroud]

➢ **Flavour tagging:**
- DL1r tagger: high-level tagger exploiting low-level ones
- The DL1r tagger output nodes are combined into a **2D b-/c-jet tagger**
- Bidimensional pseudo-continuous calibration orthogonal between b-tag and c-tag, with optimised c-tagging calibration for analysis

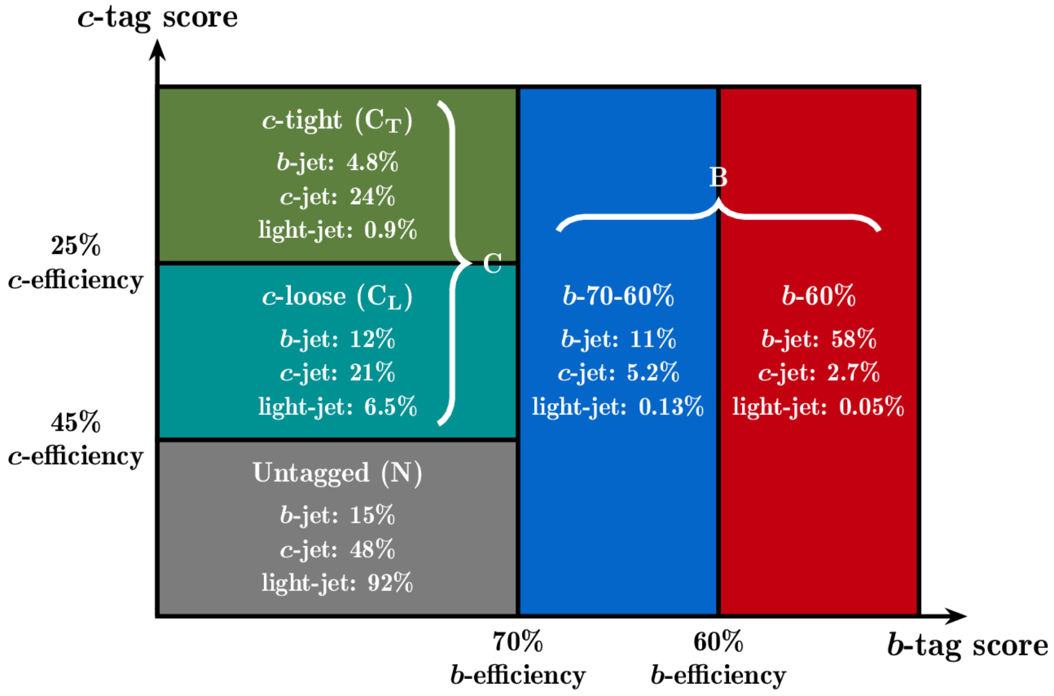


➢ **Main background**



V+jets

Top



Signal and Control Regions

Signal Regions, Total 59 (27 VH(bb), 32 VH(cc)):

For BB: Two central jets **b-tagged** (BB)
For CC: Two central jets **c-tagged** ($C_T C$)
For CC: One **c-tight tagged** jet ($C_T N$)

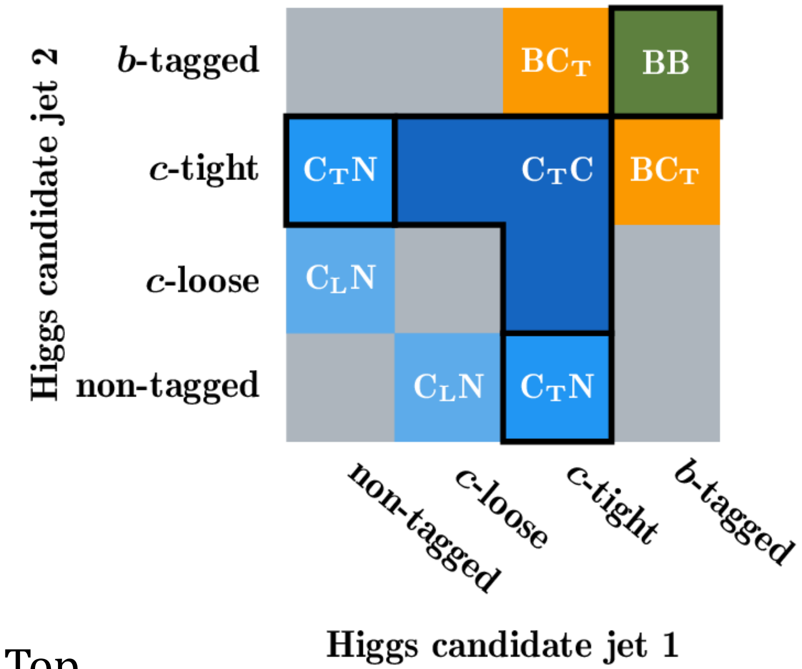
- Classified for:
  - Resolved or boosted regime
  - p_T^V
  - $\Delta R(j_1, j_2)$

Control Regions, Total 97:

The CRs are designed to target the main background V+jets and Top.
Relevant for H(cc):

1. ΔR-based CRs
- High - ΔR CRs

2. Tagging-based CRs
- $C_L N$ CRs (V+light jets)
- BC_T CRs (Top enriched)

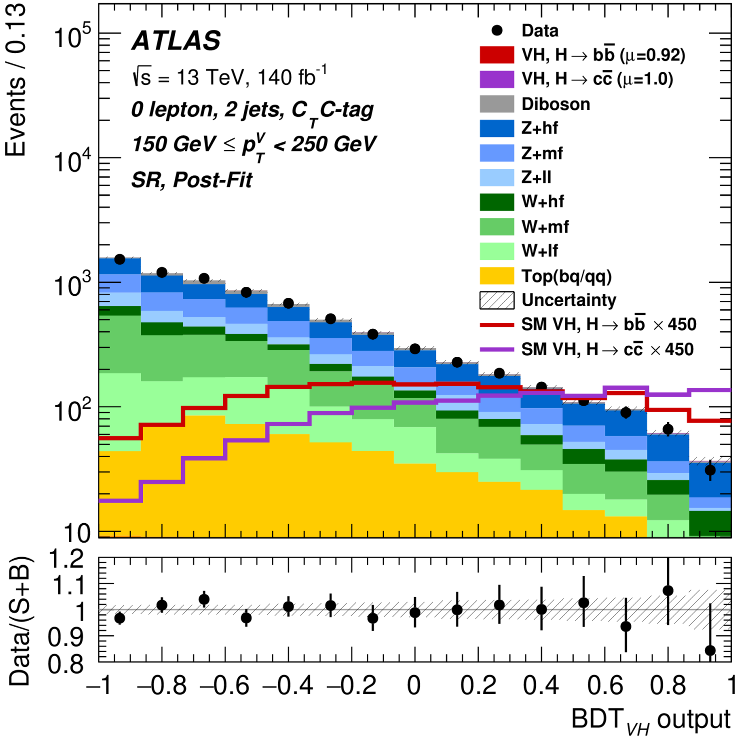


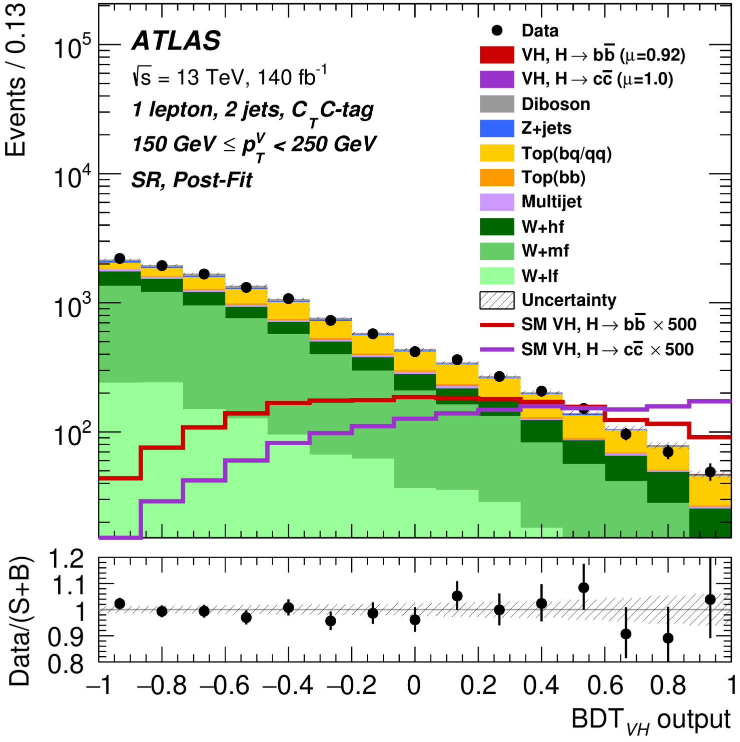
70% b-tagging efficiency

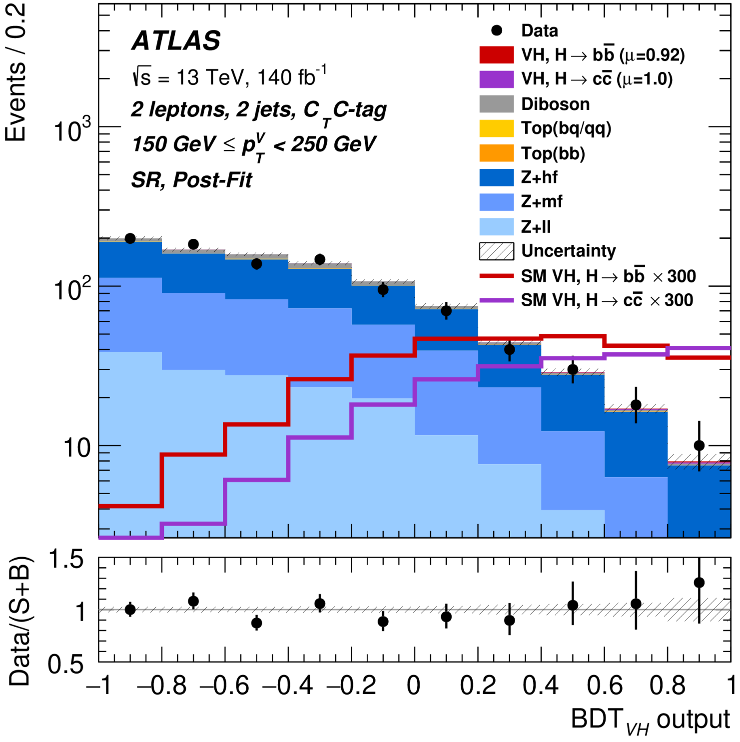
- ➢ VH(bb) enriched
- ➢ VH(cc) enriched
- ➢ Top enriched
- ➢ V + lights enriched

A Machine Learning approach

- The **Boosted Decision Tree (BDT)** is trained for each of the SR categories: BB, $C_T C_T$, $C_T C_L$ and $C_T N$ to discriminate the corresponding signal sample from the sum of the expected backgrounds.
- **In the signal regions, the BDT output is used as the likelihood binned fit variable.**
- First time in ATLAS using Machine Learning classifiers for VH(cc) and boosted VH(bb). VH(cc) significance improved by **~40% !**

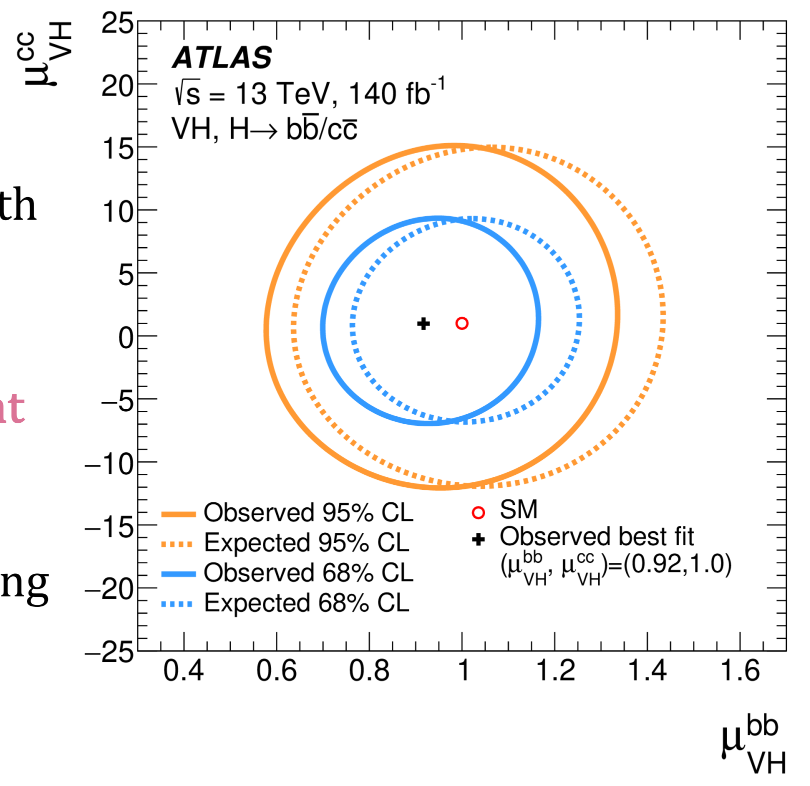


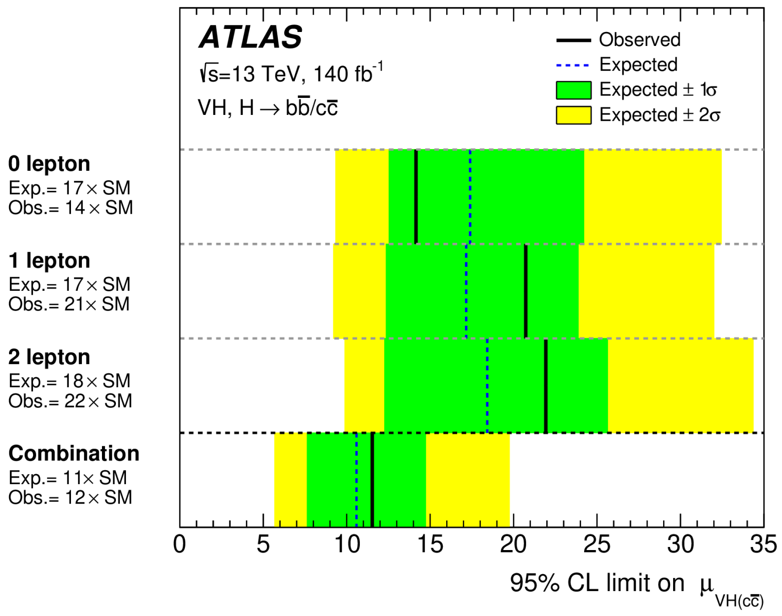


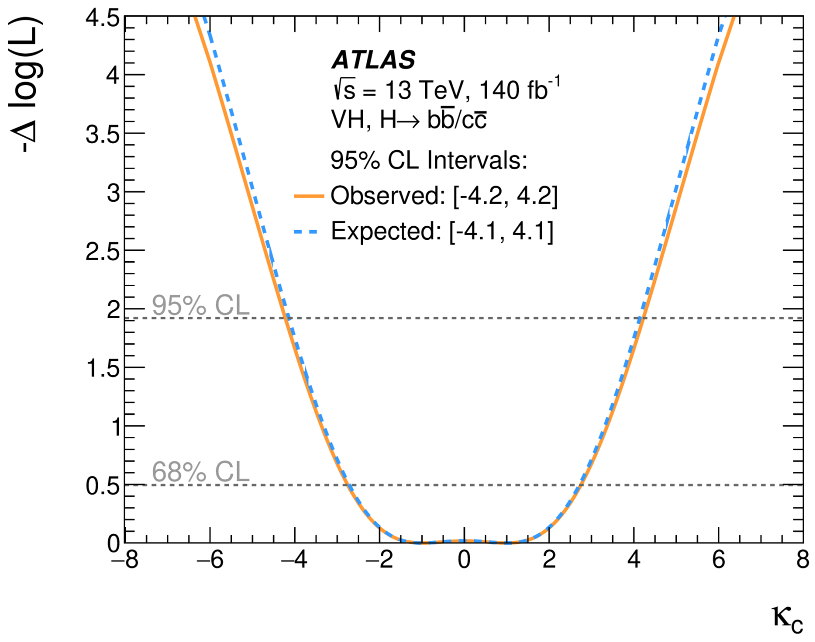


Results on VH(cc)

- The fit is always performed including all VH(bb) and VH(cc) SRs+CRs, and with the signal strengths freely floating → Signal strength $\mu = \frac{\sigma \times BR}{\sigma_{SM} \times BR_{SM}}$
- The search for the $H \rightarrow c\bar{c}$ decay yields an observed (expected) **upper limit at 95% confidence level of 11.5 (10.6)** times the Standard Model prediction.
- From the **upper limit on μ**, the charm coupling modifier is constrained resulting in **$|\kappa_c| < 4.2$ at 95% confidence level.**







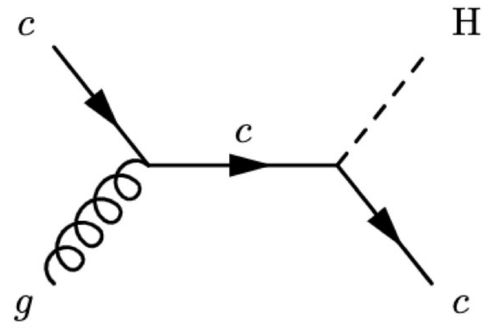
→ 3 times better than first full Run 2 results!

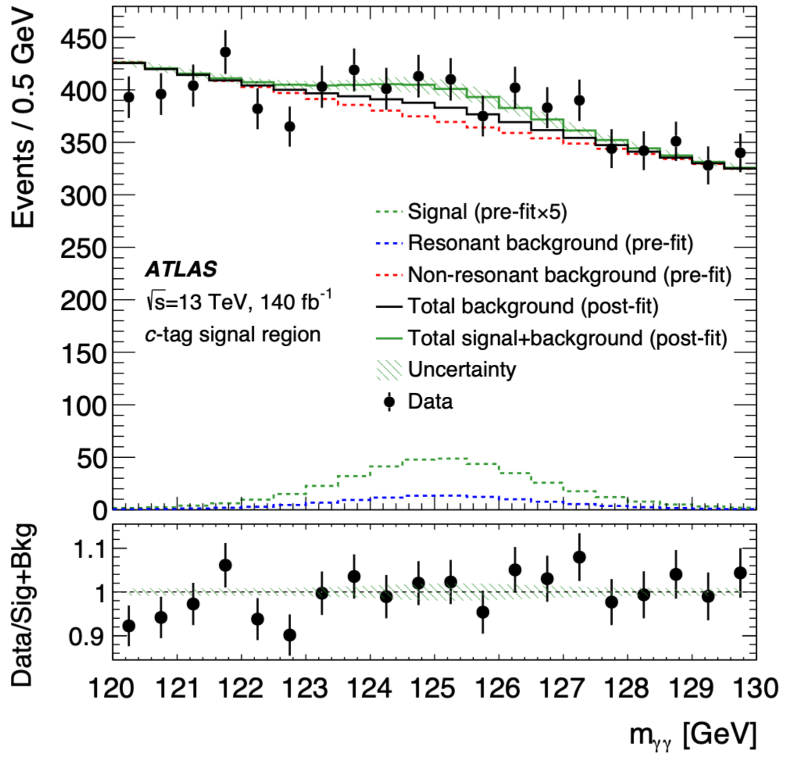
→ Mostly coming from the optimised flavour tagging scheme and BDT, also benefit from better background control!

Indirect search: H(γγ) + c production

- First direct search for **inclusive production of a Higgs boson** in association with at least one **charm jet** (H + c) exploiting the Run 2 data.

- A fraction of the H + c inclusive process is $y_c -$ sensitive, the dependency $\sigma(pp \rightarrow H + c) \propto y_c^2$ is exploited for the y_c probing!





- At least one charm-tagged jet is required in the event selection. Charm tagging is performed using DL1r.

- The background is modelled using a data-driven approach, the **Gaussian Process Regression (GPR)**.

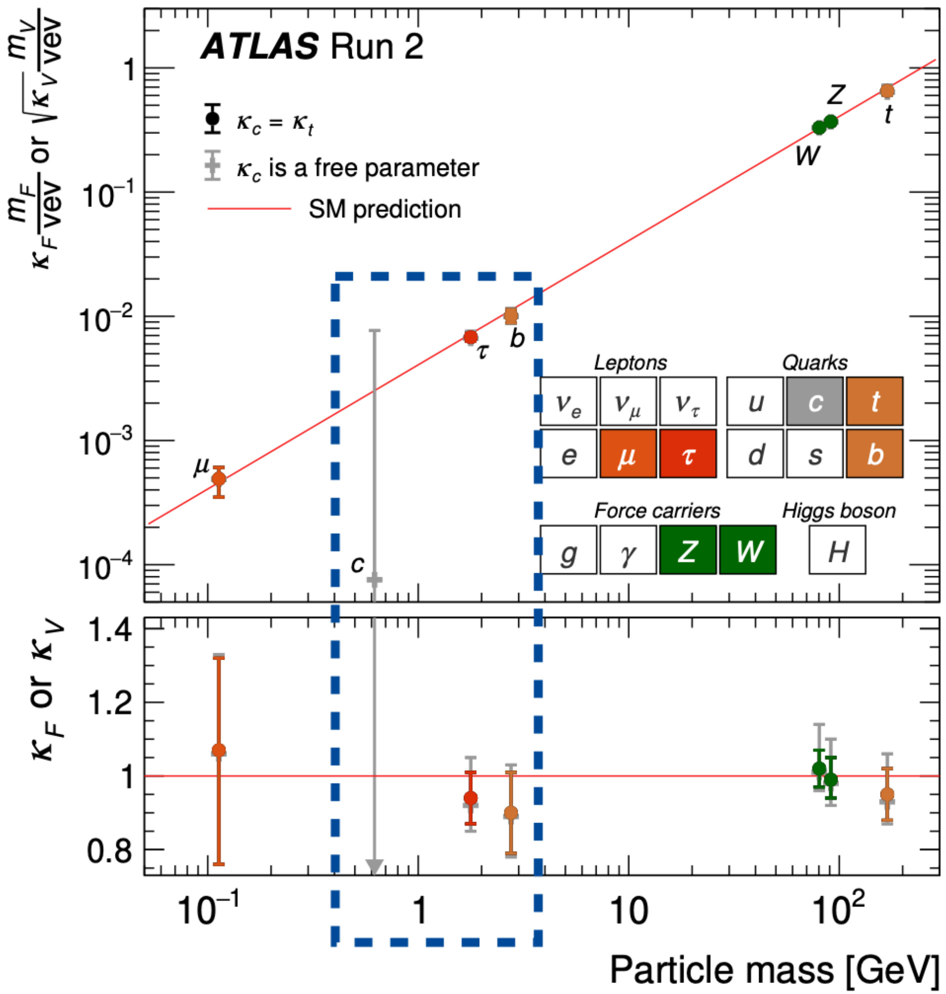
- **Best fit-value of the inclusive** H + c cross-section:

  $\sigma_{obs} = 5.3 \pm 3.2$ pb (SM expectation: 2.9 ± 3.1 pb)

- **No significant excess** over the Standard Model prediction observed, and the observed (expected) significance is 1.7σ (1.0σ)!

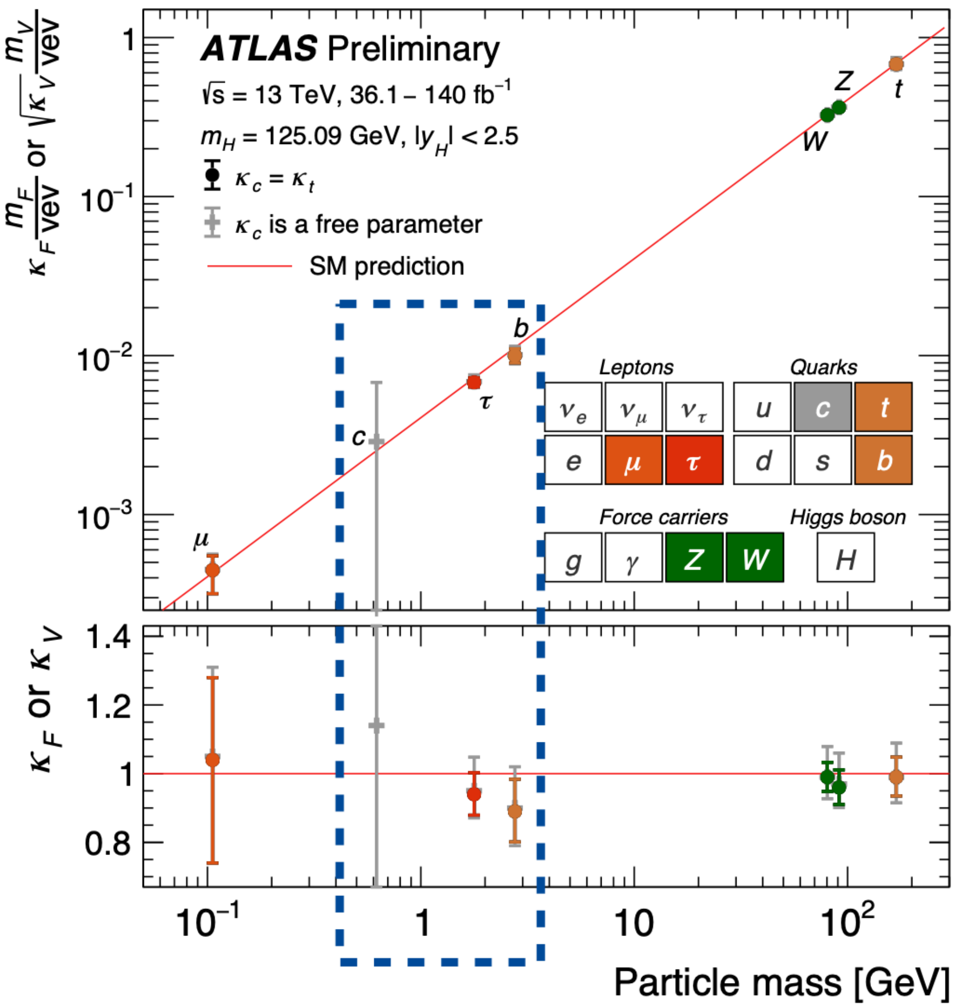
New landscape of coupling modifiers

[Nature 607 52 (2022)]



[ATLAS-CONF-2025-006]
[EPS-HEP 2025]

NEW!



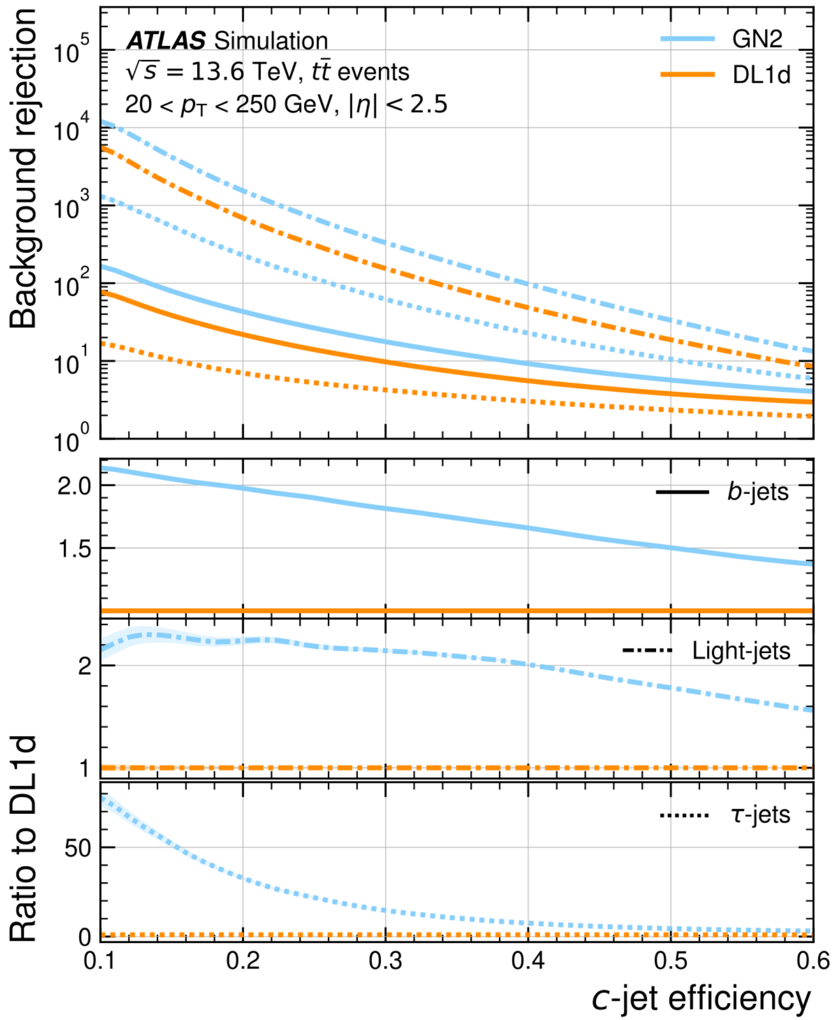
Conclusion and prospects

- Direct constrain from VH(cc) Run 2 analysis
  → $|\kappa_c| < 4.2$ @ 95% CL

- H production cross-section in association with charms probed with observed significance 1.7σ

What's next?

- Outstanding improvement in Flavour Tagging with GN2 !
- New potential channels: **VBF H(cc) and** $\mathbf{t\bar{t}H(cc)}$
- The Run 3 data set is growing with $\mathbf{> 200\ fb^{-1}}$ already recorded



Thank you for your attention!

Backup

DL1r c-jet calibration

The tagging efficiencies are measured with a data-driven method following the approach summarised here.

1. Selection of jets from $W \rightarrow cs$ in $t\bar{t}$ events

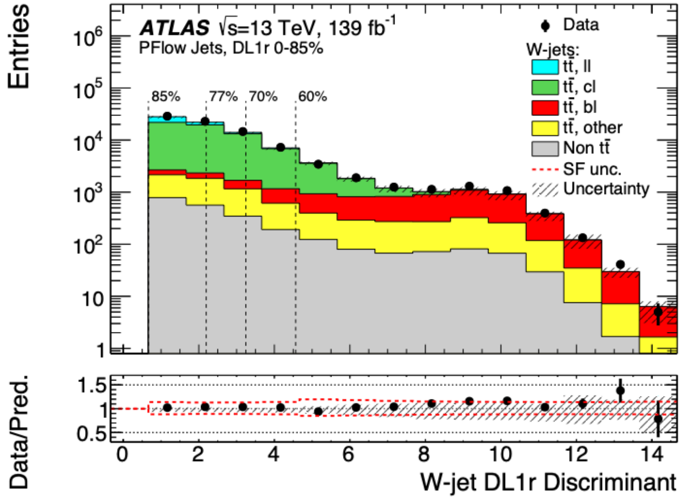
2. Measure of ϵ_{data} and ϵ_{MC} in p_T bins

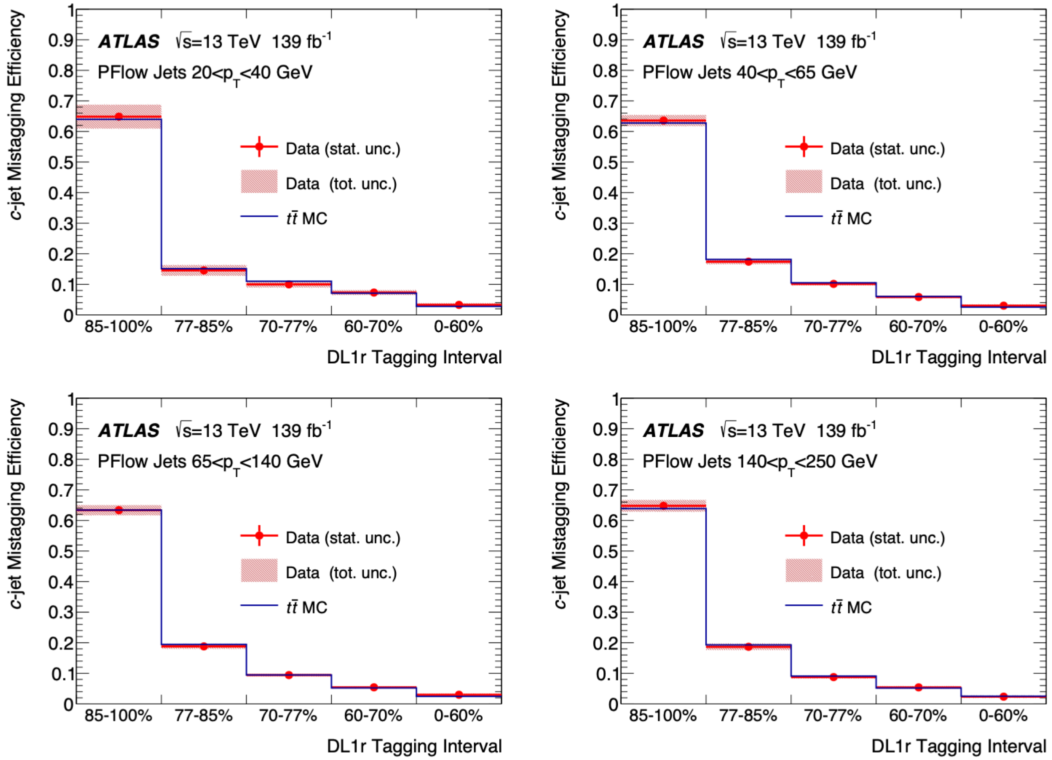
where $\epsilon = \epsilon_{mistag} = \dfrac{N_{mistagged\ c-jet}}{N_{tot\ c-jet}}$

3. Extraction of the SF:

$$SF(p_T) = \frac{\epsilon_{MC}(p_T)}{\epsilon_{data}(p_T)}$$

The results are used to correct the MC simulation.





BDT input features

Table 2: Variables used for the multivariate discriminant in each of the channels. The ✓ symbol indicates the inclusion of a variable. The $\mathrm{BDT}_{\text{Low-}\Delta R\ \text{CR}}$ uses the same variables as the 1-lepton resolved Hbb category as described in the text.

	Resolved $VH,\ H \to b\bar{b}, c\bar{c}$			Boosted $VH,\ H \to b\bar{b}$		
Variable	0-lepton	1-lepton	2-lepton	0-lepton	1-lepton	2-lepton
m_H	✓	✓	✓	✓	✓	✓
$m_{j_1 j_2 j_3}$	✓	✓	✓			
$p_{\mathrm{T}}^{j_1}$	✓	✓	✓	✓	✓	✓
$p_{\mathrm{T}}^{j_2}$	✓	✓	✓	✓	✓	✓
$p_{\mathrm{T}}^{j_3}$				✓	✓	✓
$\sum p_{\mathrm{T}}^{j_i}, i > 2$	✓	✓	✓			
$\mathrm{bin}_{D_{\mathrm{DL1r}}}(j_1)$	✓	✓	✓	✓	✓	✓
$\mathrm{bin}_{D_{\mathrm{DL1r}}}(j_2)$	✓	✓	✓	✓	✓	✓
p_{T}^{V}	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓
$E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓		✓	✓	
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$			✓			
$\lvert\Delta\phi(\boldsymbol{V},\boldsymbol{H})\rvert$	✓	✓	✓	✓	✓	✓
$\lvert\Delta y(\boldsymbol{V},\boldsymbol{H})\rvert$		✓	✓		✓	✓
$\Delta R(\boldsymbol{j_1},\boldsymbol{j_2})$	✓	✓	✓	✓	✓	✓
$\min[\Delta R(\boldsymbol{j_i}, \boldsymbol{j_1}$ or $\boldsymbol{j_2})], i > 2$	✓	✓				
N(track-jets in J)				✓	✓	✓
N(add. small-R jets)				✓	✓	✓
colour ring				✓	✓	✓
$\lvert\Delta\eta(\boldsymbol{j_1},\boldsymbol{j_2})\rvert$	✓					
$H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$	✓					
m_{T}^{W}		✓				
m_{top}		✓				
$\min[\Delta\phi(\boldsymbol{\ell}, \boldsymbol{j_1}$ or $\boldsymbol{j_2})]$		✓				
p_{T}^{ℓ}					✓	
$(p_{\mathrm{T}}^{\ell} - E_{\mathrm{T}}^{\mathrm{miss}})/p_{\mathrm{T}}^{V}$					✓	
$m_{\ell\ell}$			✓			
$\cos\theta^*(\boldsymbol{\ell^-},\boldsymbol{V})$			✓			✓

VH(bb/cc) Run 2 Analysis: Control Regions

Control Regions:
The CRs are designed to target the main background V+jets and Top.

1. ΔR-based CRs

- High - ΔR CRs (V+jets enriched): Events with larger $\Delta R(j_1, j_2)$ than SRs
- Low-ΔR CRs(Top enriched): Events with $\Delta R(j_1, j_2)$ outside the SR window

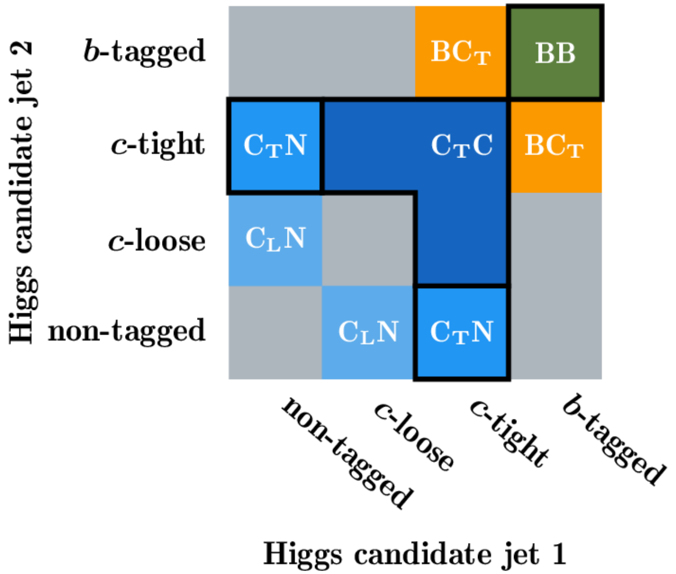
2. Tagging-based CRs

- CLN CR (V+light jets): 1 c-tagged + 1 non-tagged jet
- BCT CR (Top enriched): 1 b-tagged + 1 c-tagged jet.

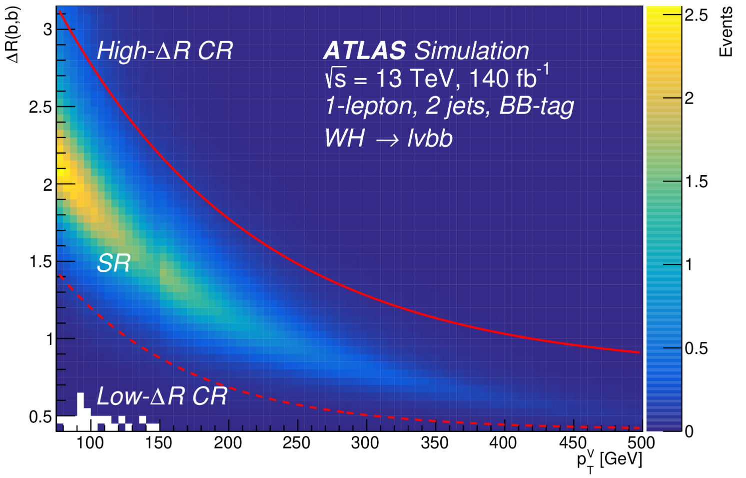
3. Leptonic-flavour CR

- eμ CR (Top enriched): electron and muon with opposite sign

+ Boosted topology CRs for H(bb) and jet multiplicity selections exploited to target different backgrounds.



- VH(bb) enriched
- VH(cc) enriched
- Top enriched
- V + lights enriched



VH(bb/cc) Run 2 Analysis: Control Regions

Control Regions:

The CRs are designed to target the main background V+jets and Top.

1. ΔR-based CRs

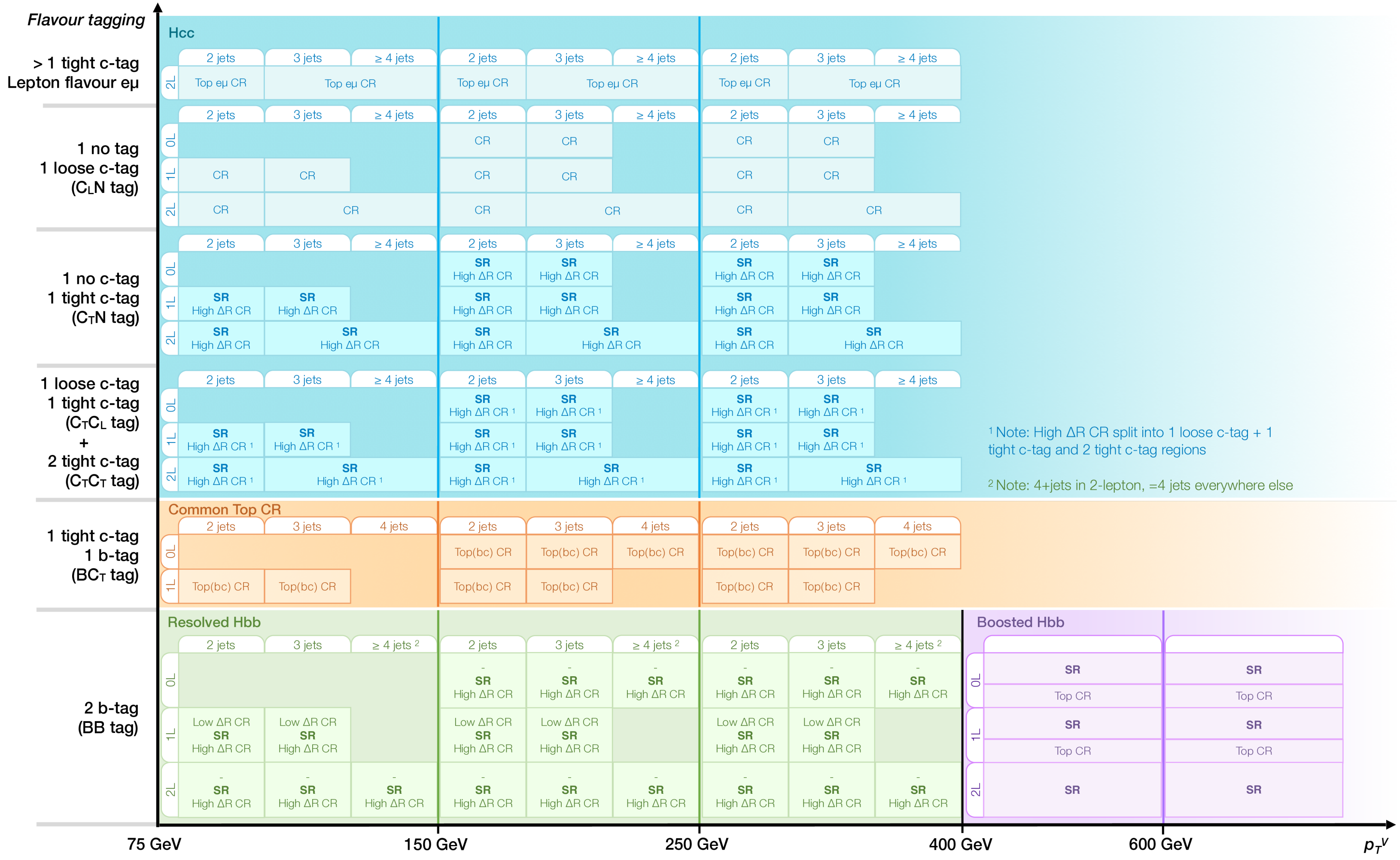
- High - ΔR CRs (V+jets enriched)
- Low-ΔR CRs(Top enriched)

2. Tagging-based CRs

- CLN CR (V+light jets)
- BCT CR (Top enriched)

3. Leptonic-flavour CR

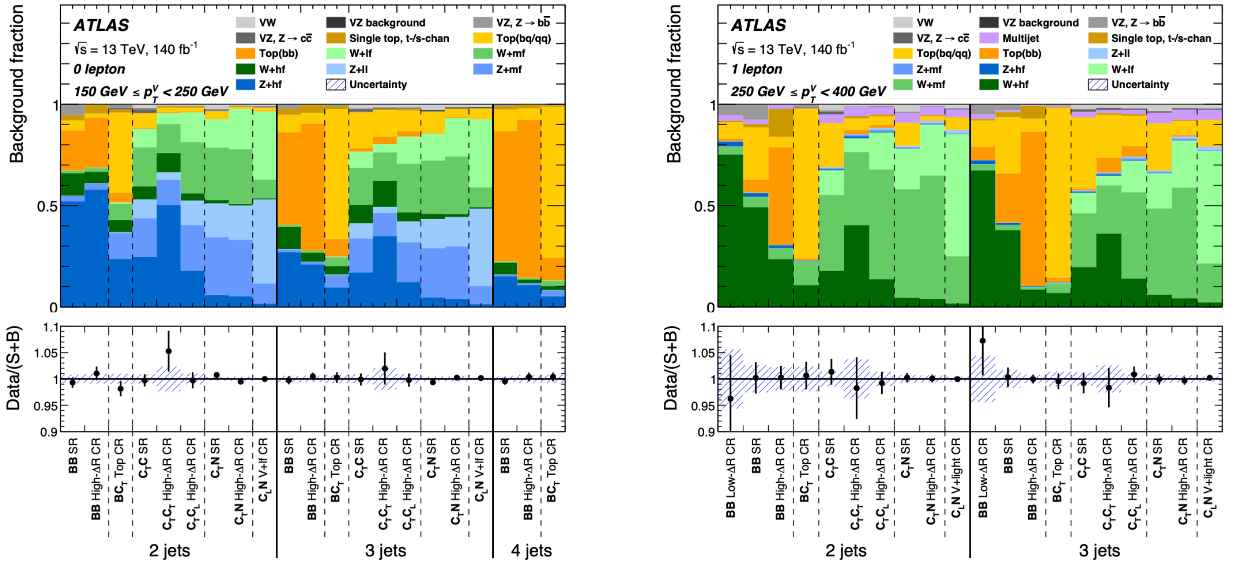
- eμ CR (Top enriched)



VH(bb/cc) Run 2 Analysis: Control Regions

Table 3: A schematic of the fit variables used in the control regions. In the signal regions, the BDT output is used as the fit variable. The 'Norm. Only' label indicates that only the event yield is used in the fits and '—' indicates that the region is not used in the fits.

Channel	Region	BB	C_TN	C_TC_L	C_TC_T	BC_T	C_LN
0-lepton	High-ΔR CR	Norm. Only				—	
	BC_T Top CR	—				$m_{j_1 j_2}$	—
	$V+lf$ CR	—					Norm. Only
1-lepton	Low-ΔR CR	$\mathrm{BDT}_{\text{Low-}\Delta R\ \text{CR}}$	—				
	High-ΔR CR	p_T^V		$m_{j_1 j_2}$		—	
	BC_T Top CR	—				$m_{j_1 j_2}$	—
	$V+lf$ CR	—					p_T^V
2-lepton	High-ΔR CR	p_T^V		$m_{j_1 j_2}$		—	
	Top $e\mu$ CR	—	Norm. Only			–	—
	$V+lf$ CR	—					p_T^V



(a) (b)

Figure 5: The relative background composition in the signal and control regions with (a) $150\,\text{GeV} < p_T^V < 250\,\text{GeV}$ in the 0-lepton channel and (b) $250\,\text{GeV} < p_T^V < 400\,\text{GeV}$ in the 1-lepton channel. The ratio of the data to the total signal-plus-background (S+B) prediction is also shown. The background predictions are adjusted with the results of the VH fit to data described in Section 9 and the uncertainty band corresponds to the overall uncertainty in the background predictions.

Uncertainty breakdown

Table 7: The breakdown of contributions to the uncertainty in the fitted value of the signal strengths (σ_μ) for $VH, H \to b\bar{b}$, $WH, H \to b\bar{b}$, $ZH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$. The sum in quadrature of uncertainties from different sources may differ from the total due to correlations. In cases where the upward and downward systematic variations have different values, the mean of the absolute values is shown.

Source of uncertainty		σ_μ			
		$VH, H \to b\bar{b}$	$WH, H \to b\bar{b}$	$ZH, H \to b\bar{b}$	$VH, H \to c\bar{c}$
Total		0.153	0.204	0.216	5.31
Statistical		0.097	0.139	0.153	3.94
Systematic		0.118	0.149	0.153	3.57
Statistical uncertainties					
Data statistical		0.090	0.129	0.139	3.67
$t\bar{t}$ $e\mu$ control region		0.009	0.014	0.027	0.08
Background floating normalisations		0.034	0.049	0.042	1.24
Other VH floating normalisation		0.007	0.018	0.014	0.33
Simulation samples size		0.023	0.033	0.030	1.62
Experimental uncertainties					
Jets		0.027	0.035	0.030	1.02
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.010	0.005	0.021	0.23
Leptons		0.003	0.002	0.010	0.25
Flavour tagging	b-jets	0.020	0.018	0.026	0.29
	c-jets	0.013	0.017	0.012	0.73
	light-flavour jets	0.005	0.008	0.008	0.66
Pile-up		0.008	0.017	0.002	0.23
Luminosity		0.006	0.007	0.006	0.08
Theoretical and modelling uncertainties					
Signal		0.076	0.074	0.101	0.72
Z + jets		0.042	0.018	0.081	1.77
W + jets		0.054	0.087	0.026	1.42
$t\bar{t}$ and Wt		0.018	0.033	0.018	1.02
Single top-quark (s-, t-ch.)		0.010	0.018	0.002	0.16
Diboson		0.033	0.039	0.049	0.52
Multijet		0.005	0.010	0.005	0.55