VH(bb) Measurement with Run2 Data

Jingyi Han
Supervisor: Lianliang Ma

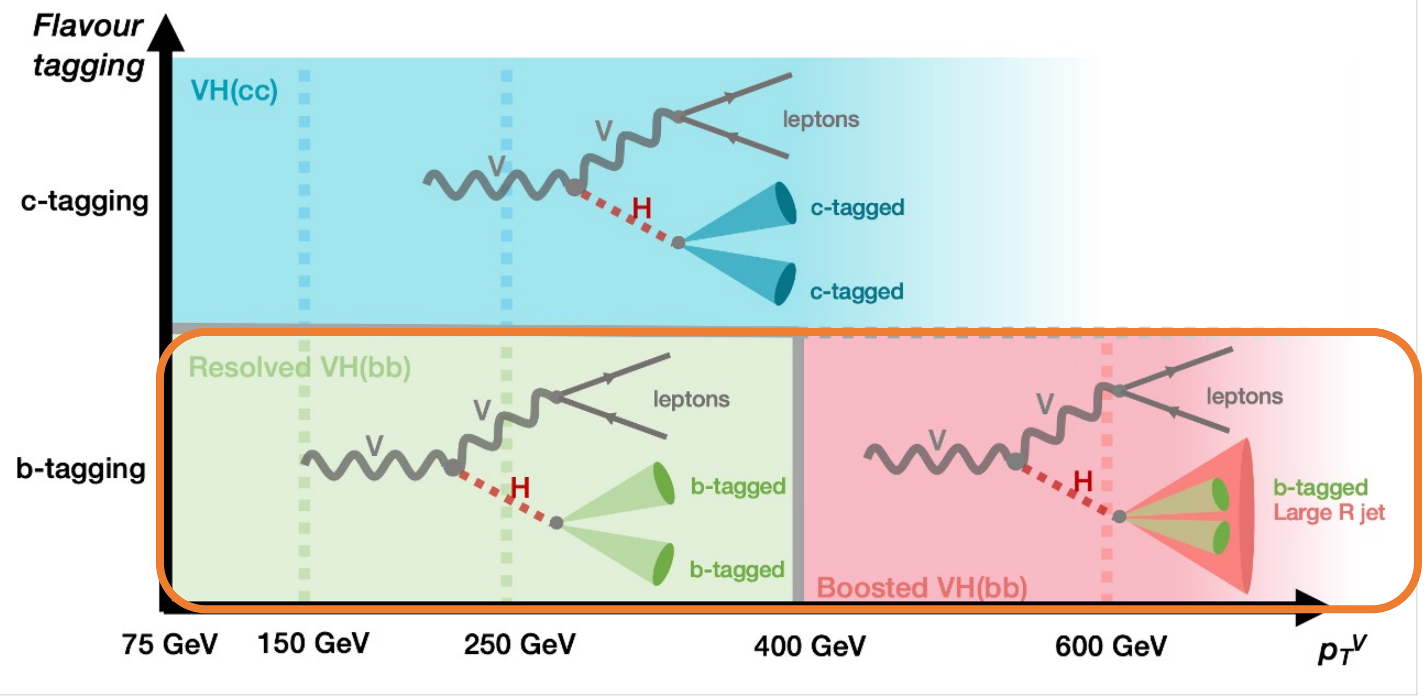
CLHCP 2023

2023.11.17

Analysis overview

VH(bb/cc) legacy analysis (INT note) is a combined measurement for **resolved VH(bb)**, **boosted VH(bb)** and **VH(cc)** analysis based on **MVA** method.

- This talk will focus on the **resolved + boosted VH(bb)** measurement



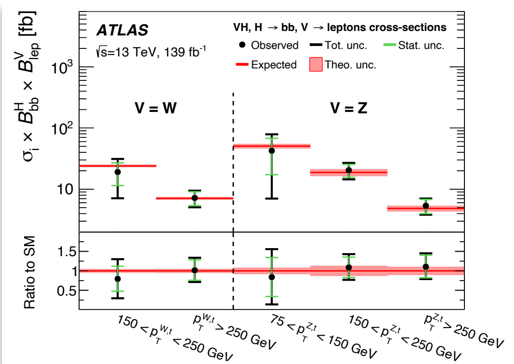
- Resolved and boosted VH(bb) split by p_T^V
- Resolved VH(bb) and VH(cc) split by **FTAG**
- Overlap between VH(cc) and boosted VH(bb) is negligible

Previous publications

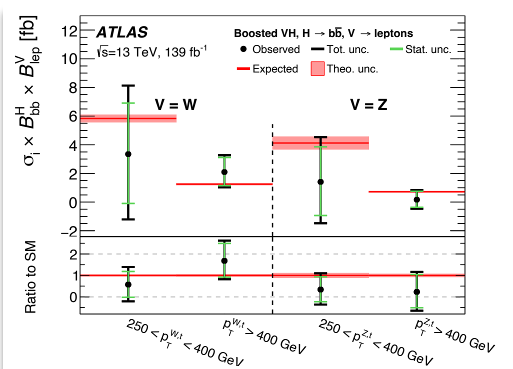
- Paper: Resolved VH(bb) analysis
- Paper: Boosted VH(bb) analysis

CONF note: resolved + boosted combination (split at 400 GeV, cut-based in boosted)

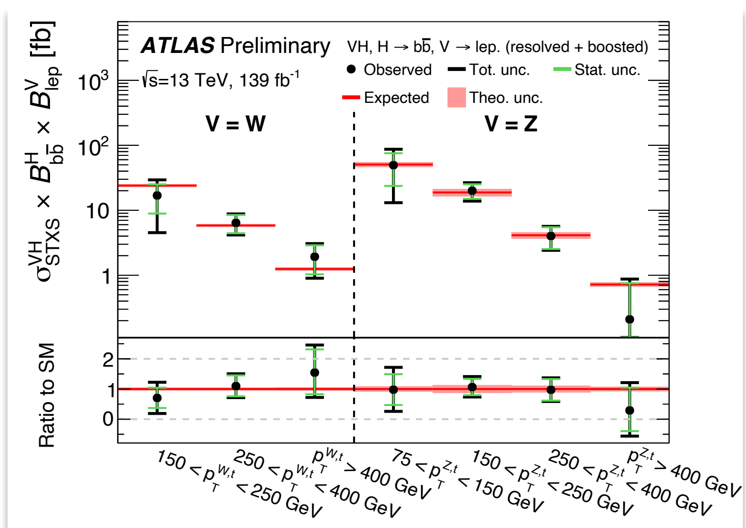
Resolved VH(bb)



Boosted VH(bb)

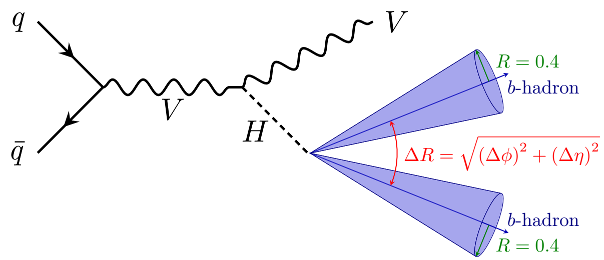


combination

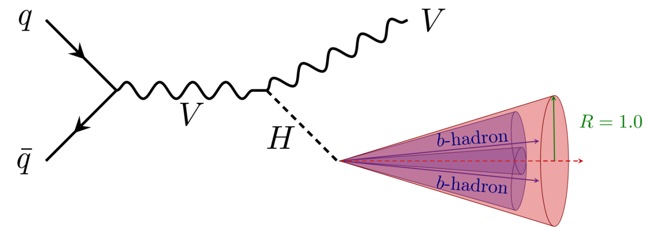


Targeted final state

resolved

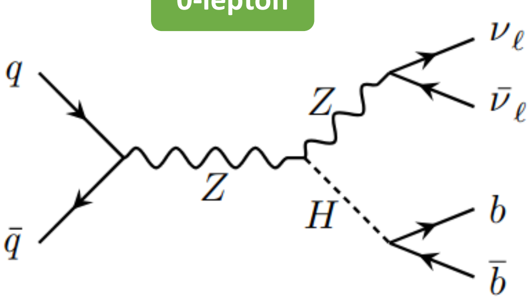


boosted

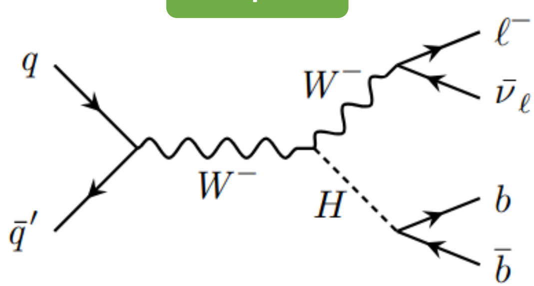


All analyses have 3 channels depending on the number of leptons in the vector boson (W, Z) decay final state: $\boldsymbol{Z \to \nu\nu}$ **(0-lepton),** $\boldsymbol{W \to l\nu}$ **(1-lepton),** $\boldsymbol{Z \to ll}$ **(2-lepton) where** $\boldsymbol{l = e, \mu}$ **(and** $\boldsymbol{\tau_{had}}$ **in 1-lepton)**

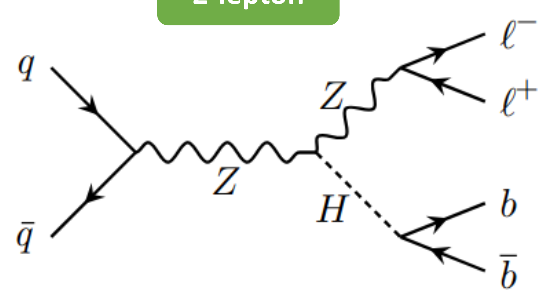
0-lepton



1-lepton

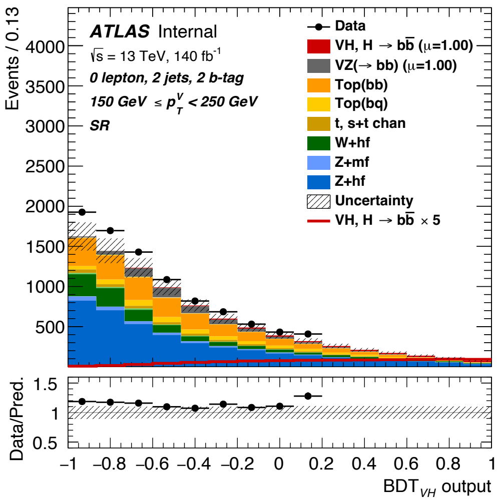


2-lepton

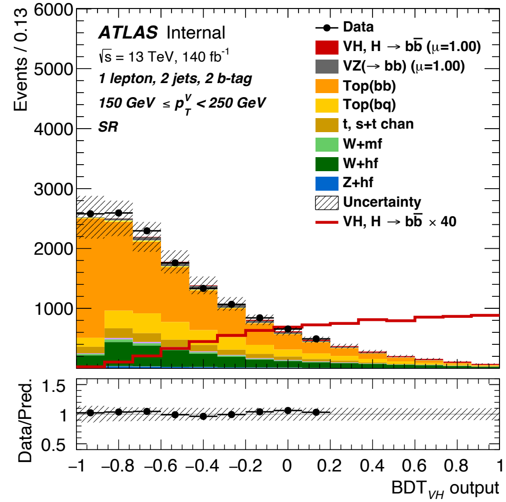


Backgrounds

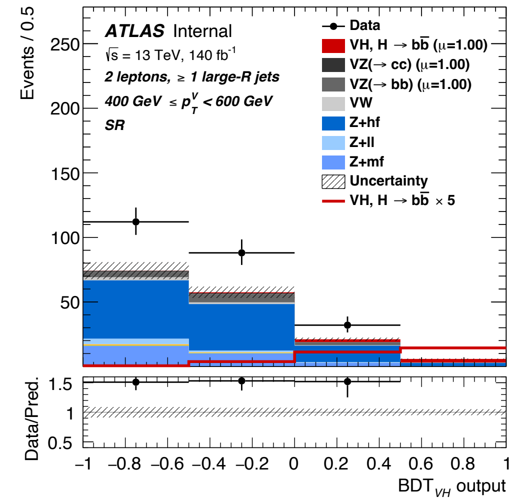
0-lepton



1-lepton



2-lepton



Main backgrounds:

- **W+jets:** W+hf (bb, cc); W+mf (bc, bl, cl); W+lf (ll) → **0L & 1L** main background
- **Z+jets:** Z+hf (bb, cc); Z+mf (bc, bl, cl); Z+lf (ll) → **0L & 2L** main background
- **Top:** ttbar and stopWt ; split into Top(bb) and other components → **0L & 1L** main background

Other small backgrounds:

- **stop:** single-top s- and t-channel
- **Diboson:** VZbb/cc is also measured in the analysis to validate the analysis strategy/ prove the robustness of the results.
- **Multi-jet** (only non-negligibel in 1L)

Object selection: leptons & MET

Electron and muon selection requirements:

- Event categorizations accroding to the number of VH-loose lepton (electron and muon)

Electron Selection	p_{T}	η	ID	d_0^{sig} w.r.t. BL	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
VH-Loose	>7 GeV	$\lvert\eta\rvert < 2.47$	LH Loose	< 5	< 0.5 mm	Loose_VarRad
ZH-Signal	>27 GeV	Same as VH-Loose				
WH-Signal	Same as ZH-Signal		LH Tight	Same as ZH-Signal		HighPtCaloOnly

Muon Selection	p_{T}	η	ID	d_0^{sig} w.r.t. BL	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
VH-Loose	>7 GeV	$\lvert\eta\rvert < 2.7$	Loose quality	< 3	< 0.5 mm	Loose_VarRad
ZH-Signal	>27 GeV	$\lvert\eta\rvert < 2.5$	Same as VH-Loose			
WH-Signal	>25 GeV when $p_{\mathrm{T}}^{V} > 150$ GeV >27 GeV when $p_{\mathrm{T}}^{V} < 150$ GeV	$\lvert\eta\rvert < 2.5$	Medium quality	< 3	< 0.5 mm	HighPtTrackOnly

tau selection requirements:

Lepton	p_{T}	η	n_{trk}	ID
tau	>20 GeV	$\lvert\eta\rvert < 2.5$	1 or 3 tracks	Loose

- Switching to **RNN ID** for hadronic tau identification. The *loose* working point of RNN ID have the same jet rejection power as BDT ID used in previous analyses, but provides around **10% higher tau efficiency.**

MET reconstructed using a track-based soft term (TST) in addition to the hard terms arising from leptons and jets.

(In this slides, the **orange color** indicates the difference/improvemennt compared to the previous analysis)

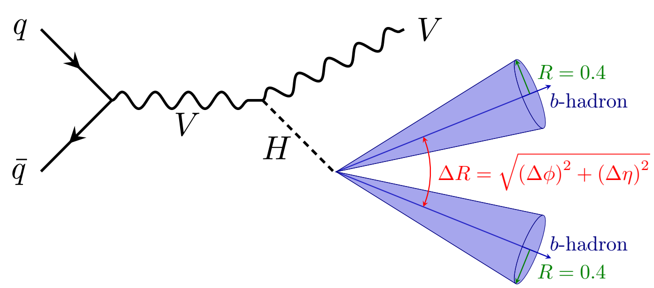
Object selection: Small-R jets

Small-R jets: calorimeter jet collection with R=0.4

- Switched to **PFlow** small-R jets, previously used EMTopo jets
  - **1-3%** improvement on di-jet mass resolution
- Used to reconstruct the Higgs candidate for VH(cc) and resoved VH(bb), and used in the event categorisation to improve the signal sensitivity for boosted VH(bb)
- Small-R jets selection requirements:
  - Tight jet cleaning is used, loose WP used previously
  - Leading signal jet $p_T > 45$ GeV

Jet Category	Tight Jet Cleaning	p_T	η	JVT/fJVT
Signal Jet	true	> 20 GeV	< 2.5	JVT>0.5 for $\lvert\eta\rvert$ <2.4, $p_\mathrm{T} < 60$ GeV jets
Forward Jet	true	> 30 GeV	$2.5 < \lvert\eta\rvert < 4.5$	fJVT<0.5 for $p_\mathrm{T} < 120$ GeV jets

- Loose fJVT now applied by default for pile-up rejection, not used in previous analyses

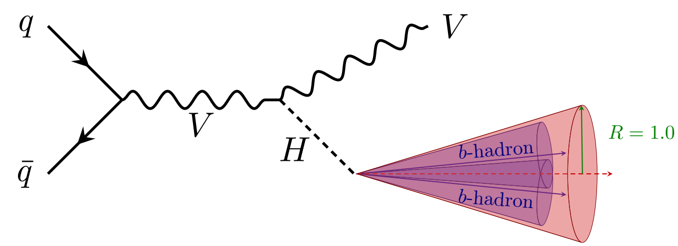


Object selection: Large-R jets and track-jets

Large-R jets: calorimeter jet collection with R=1.0

- Used to reconstruct the Higgs candidate for boosted VH(bb)
- Large-R jets selection requirements:

Jet Category	Selection Requirements
Large-R jets	trimmed AntiKt10LCTopoTrimmedPtFrac5SmallR20Jets
	$p_{\mathrm{T}} > 250$ GeV
	$\lvert\eta\rvert < 2.0$



- Large-R jet with in-situ calibration

Track-jets: AntiKtVR30Rmax4Rmin02TrackJets

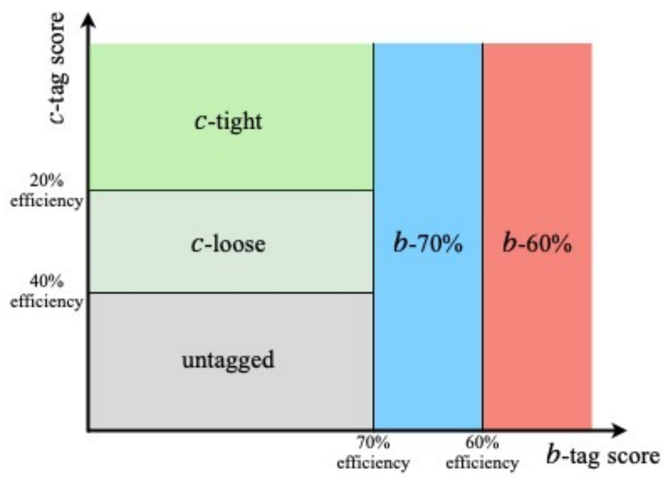
- Used to reconstruct the b-tagged objects inside the large-R jet and in the event categorization to construct a top enriched control region.
- Track-jets selections: $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.5$

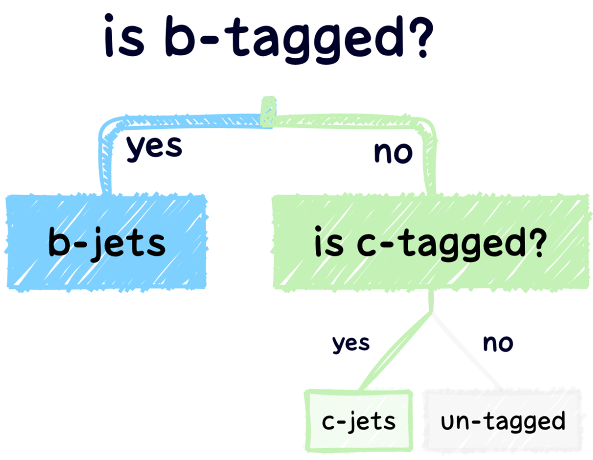
(The standard object **overlap removal** procedure is used in this analysis.)

Flavour tagging overview

Resolved VH(bb) and VH(cc) analysis: pseudo-continuous flavour tagging scheme (PCFT)

- b-tagging WPs: DL1r 60%, 70%
- c-tagging WPs: DL1r tight, loose WPs





Boosted VH(bb):

- b-tagging WP: DL1r 85%
- Xbb tagger shows very similar performance with DL1r
- Corss checked using the Xbb tagger calibration strategy, that the the close-by effects of DL1r are negligible.

Event Selection

Resolved VH(bb) & VH(cc)

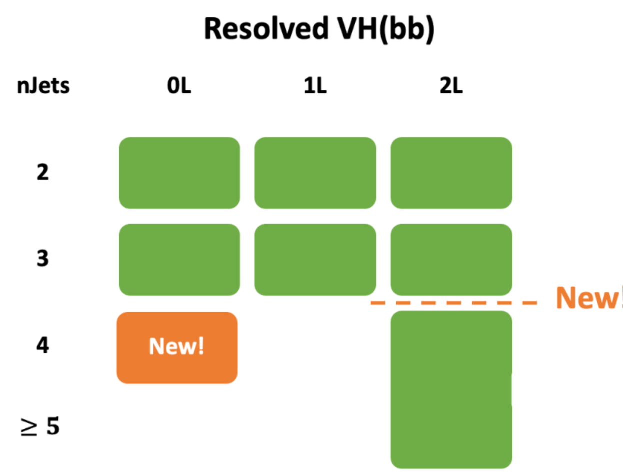
Analysis regime	$VH(\to b\bar{b})$	$VH(\to c\bar{c})$
	Common Selections	
Jets	≥ 2 signal jets	
Candidate jets tagging	2 B-tags	≥ 1 T-tag [a]
Leading Higgs candidate jet p_T	> 45 GeV	
Sub-leading Higgs candidate jet p_T	> 20 GeV	
m_{bb} or m_{cc}	> 50 GeV (before correction)	
ΔR(jet1, jet2)	Upper cut $\Delta R < \pi$	
	0 Lepton	
Trigger	E_T^{miss} triggers	
Jets	≤ 4 jets	≤ 3 jets
Additional jets tagging	no T-tag	no B-tag[b]
Leptons	0 VH-loose lepton	
E_T^{miss}	> 150 GeV	
S_T	> 120 (2 jets), >150 GeV (3p jets)	
m_T^W	> 10 GeV (for events with at least one hadronic τ)	
$\lvert\min\Delta\phi(E_T^{\text{miss}}, \text{jet})\rvert$	$> 20°$ (2 jets), $> 30°$(3 jets)	
$\lvert\Delta\phi(E_T^{\text{miss}}, H)\rvert$	$> 120°$	
$\lvert\Delta\phi(\text{jet1}, \text{jet2})\rvert$	$< 140°$	
	1 Lepton	
Trigger	e channel: single electron trigger μ channel: single muon trigger ($p_T^V < 150$ GeV) and E_T^{miss} triggers (above)	
Jets	≤ 3 jets	
Additional jets tagging	no T-tag	no B-tag??
hadronic τ-veto	no hadronic τ	
Leptons	1 WH-signal lepton > 1 VH-loose lepton veto	
E_T^{miss}	> 30 GeV (e channel)	
S_T	> 120 (2 jets), >150 GeV (3 jets) (μ channel with E_T^{miss}trigger)	
m_T^W	> 20 GeV ($75 < p_T^V < 150$ GeV only)	
	2 Lepton	
Trigger	e channel: single electron trigger μ channel: single muon trigger ($p_T^V < 250$ GeV) and E_T^{miss} triggers (above)	
Additional jets tagging	-	no B-tag
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton) Same flavour[c], opposite-charge for $\mu\mu$	
m_{ll}	$81 < m_{ll} < 101$ GeV	

[a] Except for $V + lights$ CRs, where exactly one L-tag jet and no T-tag jet is required.
[b] Except for Top BT CR, where the ≥ 1 B-Tag jets is required
[c] Except for Top $e\mu$ CR, where different flavor is required

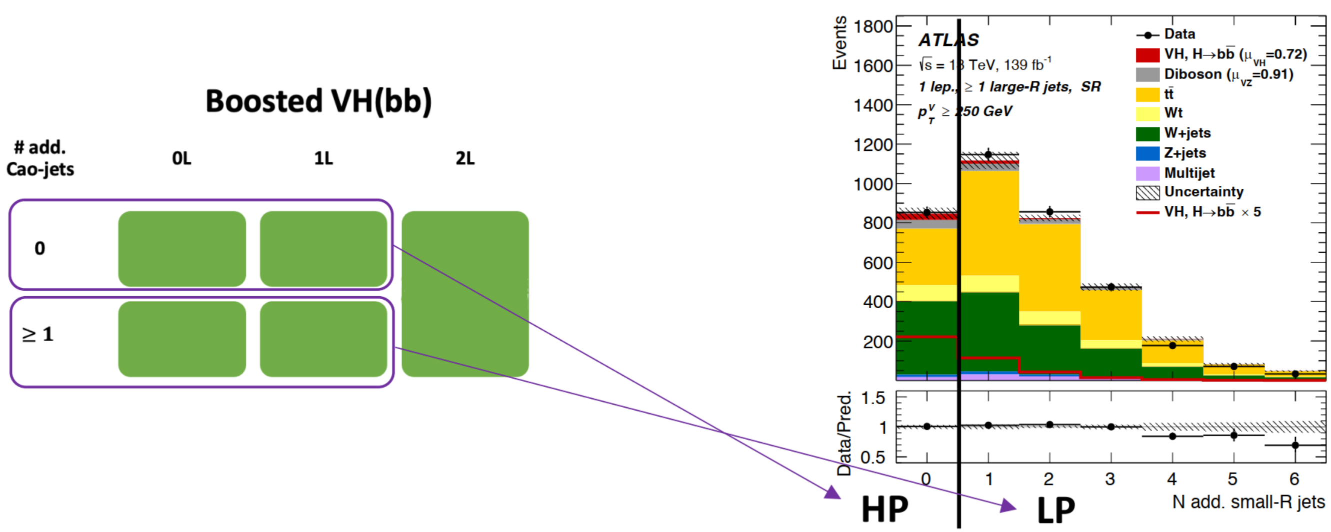
Boosted VH(bb)

Selection	0 lepton channel	1 lepton channel		2 lepton channel	
		e sub-channel	μ sub-channel	e sub-channel	μ sub-channel
Trigger	E_T^{miss}	Single electron	E_T^{miss}	Single electron	E_T^{miss}
Leptons	0 VH-*loose* lepton	1 WH-*signal* lepton no second VH-*loose* lepton no hadronic τ		≥ 1 ZH-*signal* lepton 2 VH-*loose* leptons	
E_T^{miss}	> 250 GeV	> 50GeV	-	-	
p_T^V	$p_T^V > 400$ GeV				
Large-R jet	at least one large-R jet, $p_T > 250$ GeV, $\lvert\eta\rvert < 2$				
Track-Jets	at least two track-jets, $p_T > 10$ 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				
m_J	> 50 GeV				
min $\Delta\phi(E_T^{\text{miss}}, \text{jets})$	$> 30°$	-			
$\Delta\phi(E_T^{\text{miss}}, H_{\text{cand}})$	$> 120°$	-			
$m_{\ell\ell}$	-			66 GeV$< m_{\ell\ell} <$ 116 GeV	
lepton flavor	-			two lepton same flavour	
lepton charge	-			opposite sign muons	

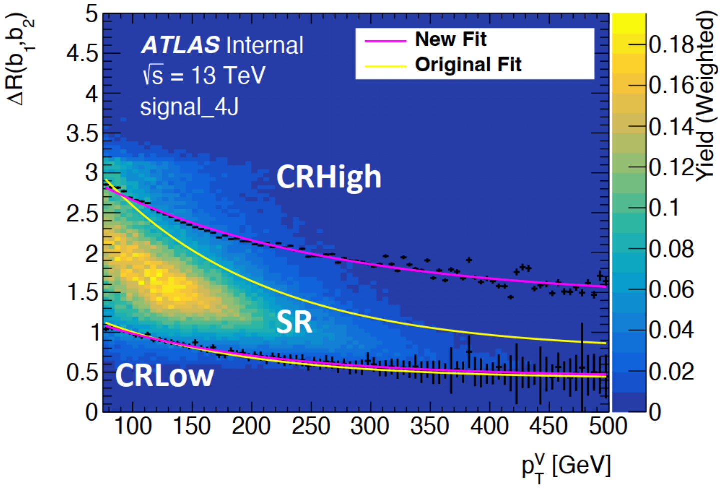
Event categorization: pTV & nJet



(Merge 0L $n_{\tau_{had}} \geq 1$ events into 1L)



Event categorization: ΔR control region



Defined by continuous $\Delta R(b_1 b_2)$ - p_T^V cut.

The cut is optimized seperatly for 2J, 3J, 4J and 5pJ categories

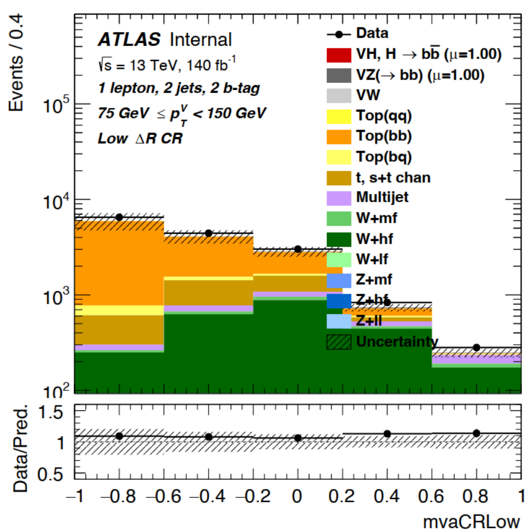
High ΔR Control Region (CRHigh):

- Designed to control Z+jets (0/2L) and Top(0L/1L)

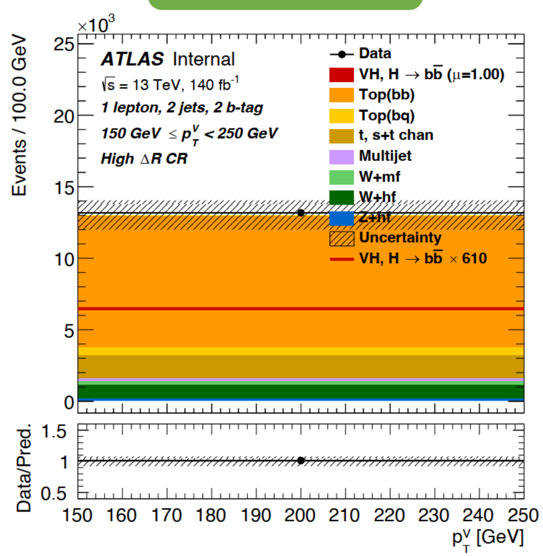
Low ΔR Control Region (CRLow):

- Only used in resolved VH(bb) 1L channel to constrain W+hf.

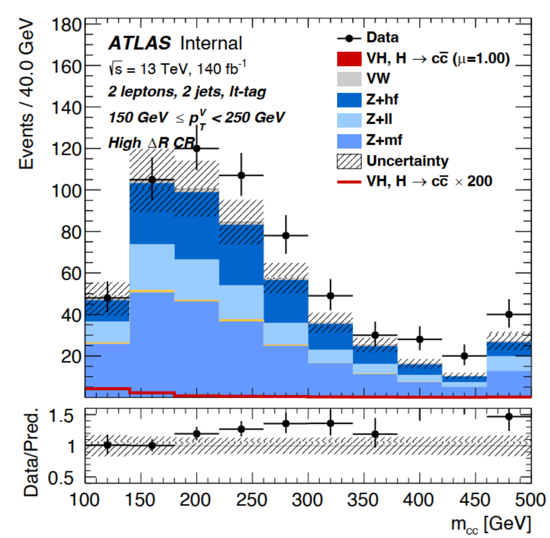
1L CRLow



1L CRHigh



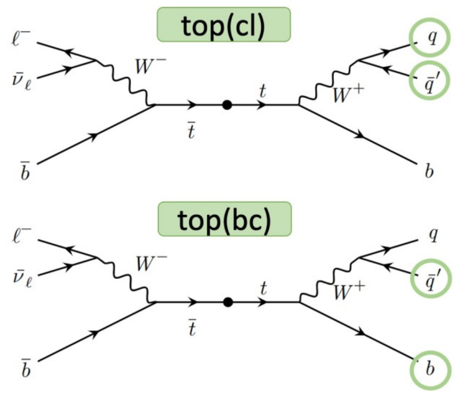
2L CRHigh

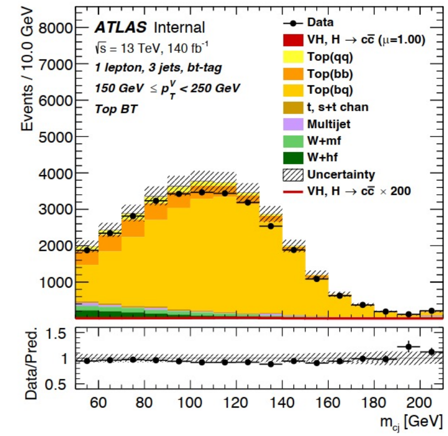


Control regions: top control region

Resolved top(bc) CR:

- $\geq$ 1 b-tagged jets and $\geq$ 1 c-tagged jets
- Used in **0L/1L**: much better constraint on Top(bc); **CRHigh** is enriched with **Top(bb)**

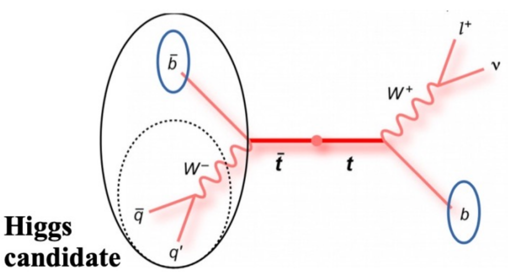


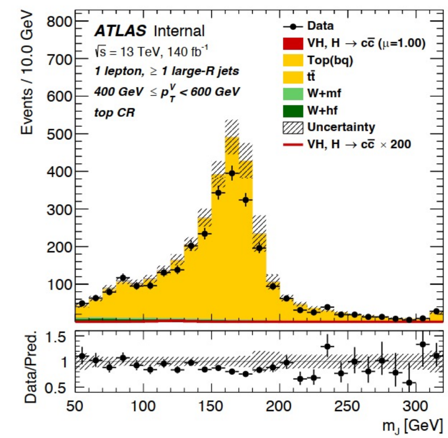


Boosted top CR:

- at least one additional b-tagged track jet outside the Higgs-candidate jet.
- b-tagging WP: DL1r@85%

  85% WP implies that a lot of ttbar events moved from SR to topCR, **improves purity in topCR**





Modelling: MC samples

sample	Nominal	Alternative
VH(bb)	PwPy8	Powheg+Herwig7
Diboson	Sh2.2.11	PwPy8 Sh2.2.1
ttbar	PwPy8	Powheg+Herwig7 aMC@NLO+Pythia8(DR)
Single-top	PwPy8(DS)	Powheg+Herwig7 PwPy8(DR) aMC@NLO+Pythia8(DR)
V+jets	Sh2.2.11	MG FxFx

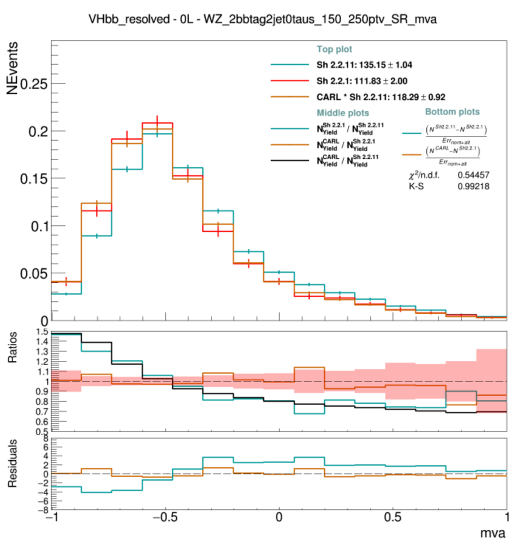
- 1L multi-jet and 2L top background are estimated by data-driven

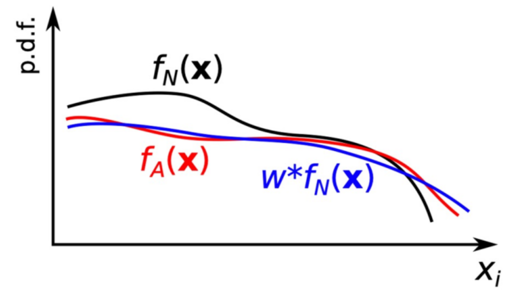
Modelling: strategy

The modelling uncertainties are derived separately for each process and are divided into **normalisation/acceptance uncertainties and shape uncertainties.**

Shape uncertainties for most sample are calculated by **CARL rewighting:**

- **reweighting:** motivated by low statistics of alternative sample. Take events sampled from the nominal and re-weight them event-by-event to make the re-weighted distribution looks like the alternative.
- **CARL** is a re-weighting technique based on **DNN.** Simultaneously reweight multiple variables, and **correlations** between them are considered, important as now **SRs are fitted with BDT**





Top panel:

- Green: Nominal (Sh2.2.11) sample
- Red: Alternative (Sh2.2.1) sample
- Orange: reweighted nominal

Middle panel: ratios

- Green (Black): ratio of alternative (Nominal * CARL) over the nominal
- Orange: ratio of Nominal * CARL over the alternative sample --> **should be at 1**

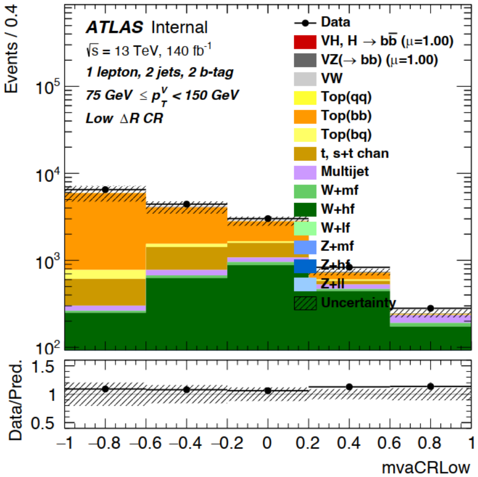
Bottom panel: residual = (ratio - 1)/relative error

- Green: residual of nominal
- Orange: residual of Nominal*CARL --> **should be at 0**

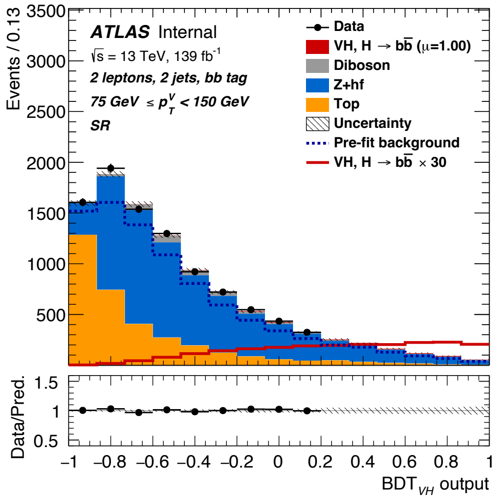
Fit overview

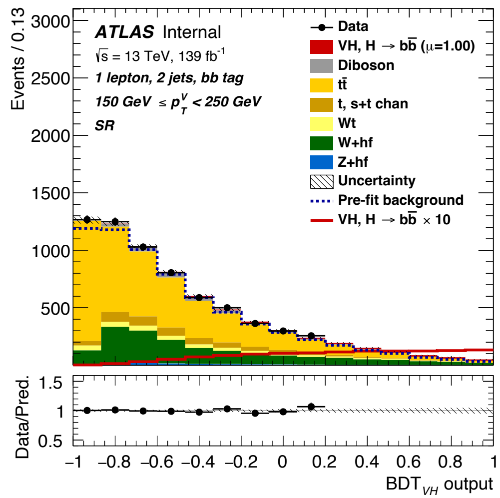
- A simultaneously fit is performed for resolved VH(bb), boosted VH(bb) and VH(cc) in the legacy analysis
  - Only **combined VH(bb)** fit result is shown in this slides
- SRs will be fitted with **BDT score**
- VH(bb) CRs in the fit and the variables fitted in each region:
  - Better constrain of W+hf in CRLow fitted to mvaCRLow

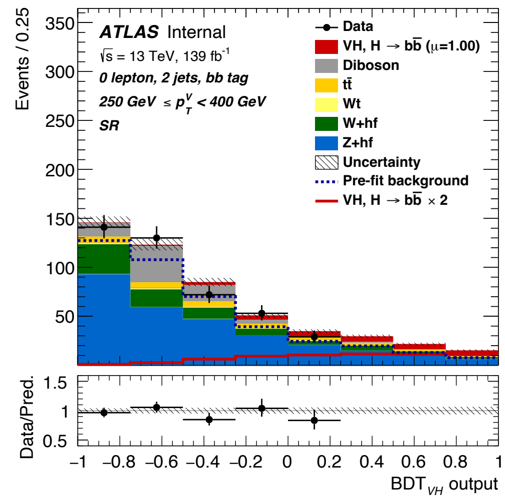
channel	region	variable
0-lepton	CRHigh	*Norm-only*
	Top(bc) CR	*mBB*
1-lepton	CRLow	*mvaCRLow*
	CRHigh	*Norm-only*
	Top(bc) CR	*mBB*
2-lepton	CRHigh	*pTV*

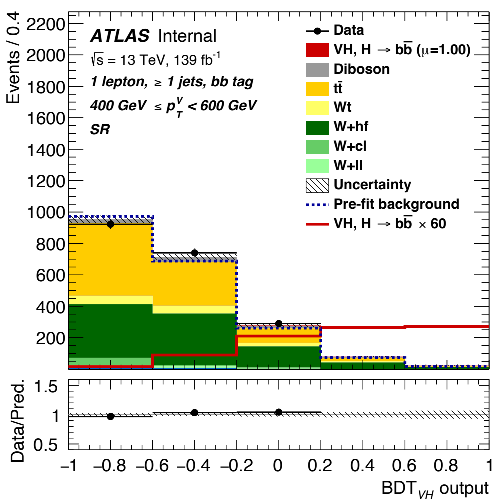


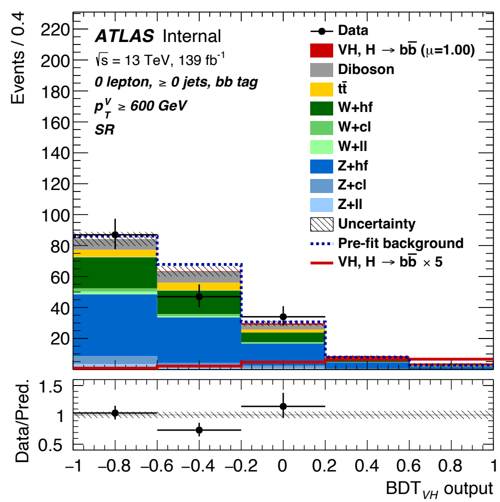
Postfit plots





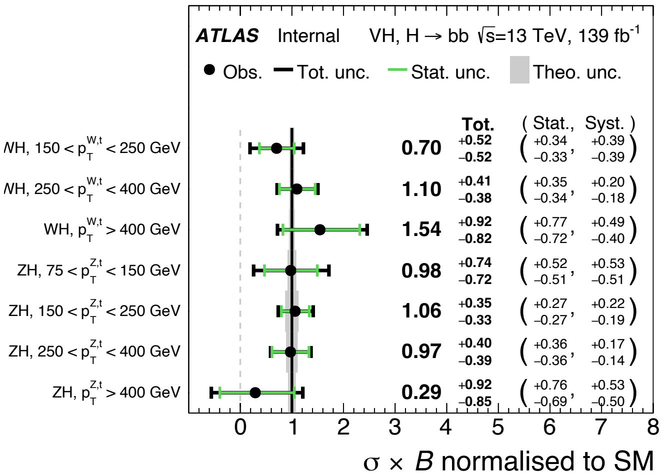




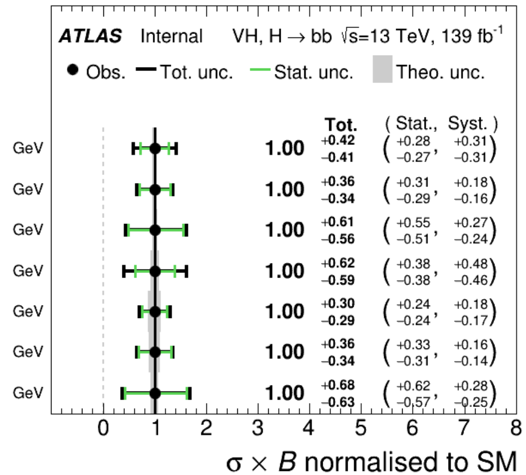


Expected results

Combination CONF note



VH(bb/cc) expect



Expected reduction of Syst.

WH, 150 < $p_T^{W,t}$ < 250 GeV	-19%
WH, 250 < $p_T^{W,t}$ < 400 GeV	-12%
WH, $p_T^{W,t}$ > 400 GeV	-34%
ZH, 75 < $p_T^{Z,t}$ < 150 GeV	-16%
ZH, 150 < $p_T^{Z,t}$ < 250 GeV	-14%
ZH, 250 < $p_T^{Z,t}$ < 400 GeV	-10%
ZH, $p_T^{Z,t}$ > 400 GeV	-26%

Summary

- Today's talk covered physics objects, event selection and categorisation, MVA, general background modelling strategy and an quick overview of the fit
- The second unblinding approval talk is coming soon.
  - A more detailed discussion about background modelling , fit model and expected results

Backup

Main deliverables

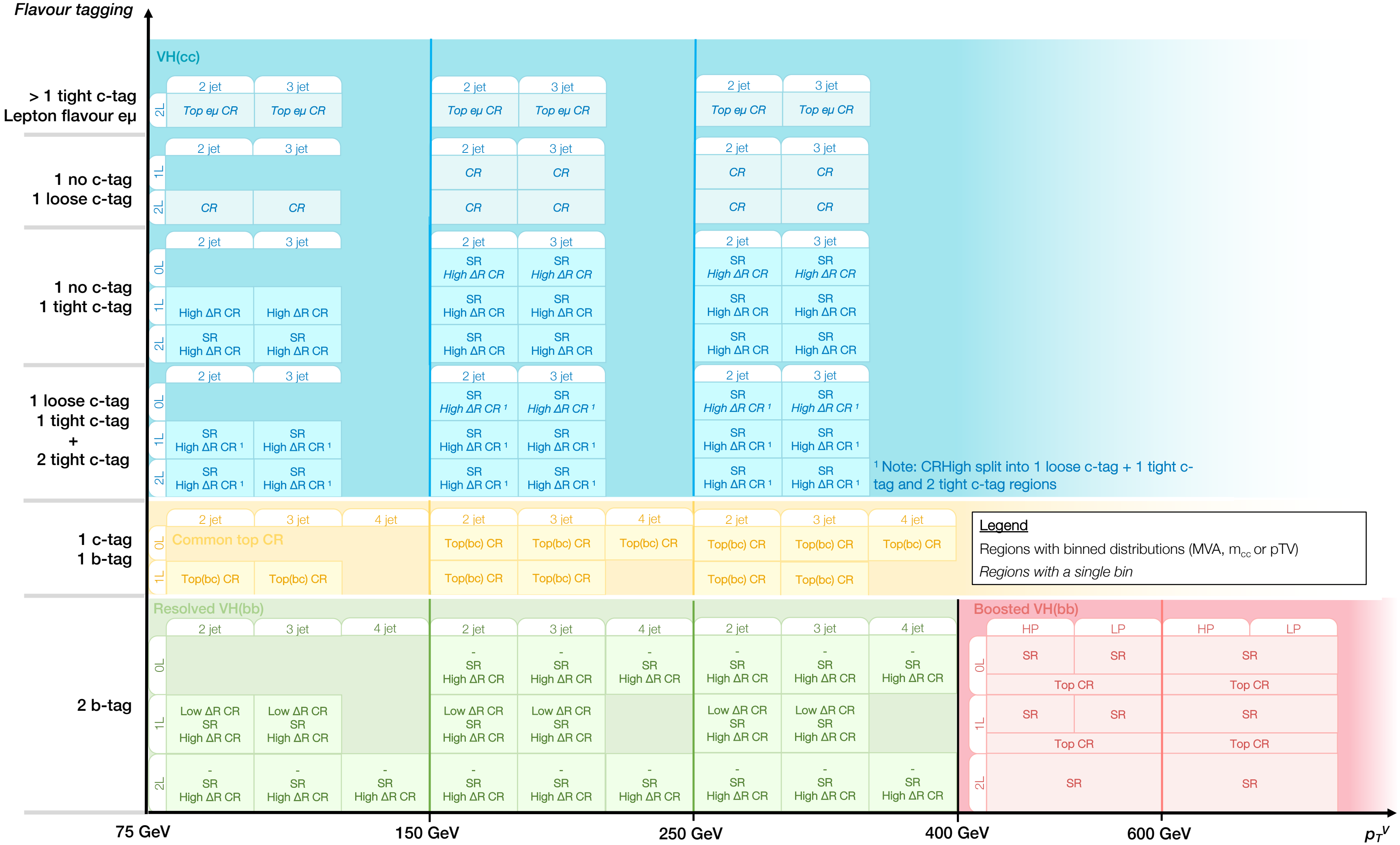
The expected deliverables for the VH(bb) process are:

- Inclusive signal strength (μ_{VH}^{bb}) and significance
- Signal strengths for WH(bb) / ZH(bb) signals (μ_{ZH}^{bb}, μ_{WH}^{bb})
- Fiducial STXS measurement in different pTV & nJ bins
- Constraint on the Higgs-bottom coupling modifier κ_b
- Effective Field Theory (EFT) interpretation for VH(bb)

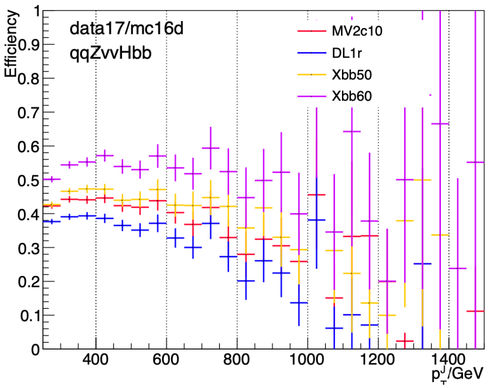
The expected deliverables for the VH(cc) process：

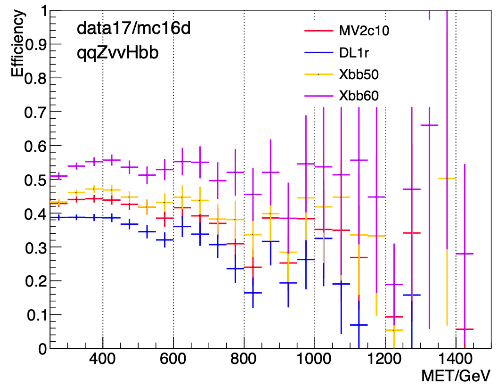
- Signal strength of the VH(cc) signal (μ_{VH}^{cc})
- Expected and observed upper limit (at the 95% CL) on VH(cc) process
- Constraint on the Higgs-charm coupling modifier κ_c and the ratio κ_c/κ_b

Analysis regions overview

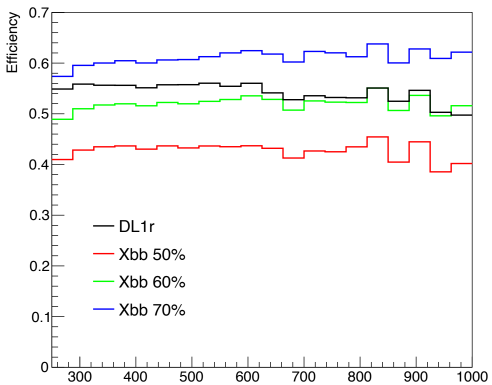


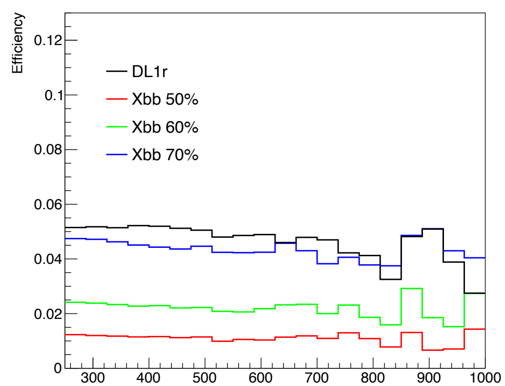
Xbb performance: DL1r vs. Xbb (link)



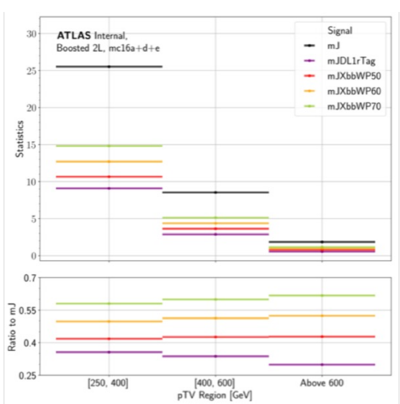


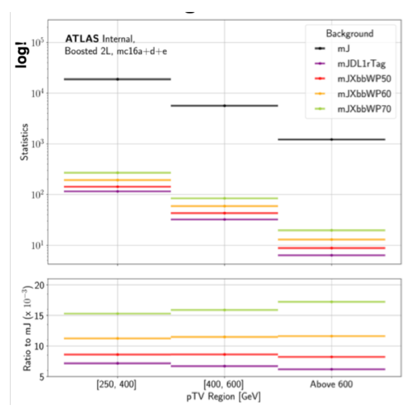
0L signal effieiency
(DL1r/MV2c10 @ 70%WP)





1L signal (left) and ttbar (right) effieiency
(DL1r @ 85%WP)





2L signal (left) and ttbar (right) yields
(DL1r @ 70%WP)

Xbb performance: DL1r vs. Xbb

DL1r@85% VS. Xbb **[MVA]**

	MVA significance	**comments**
0L	Similar for 50% and 60%	70% WP a bit worse (~3%) inn 400_600 GeV ~10% wores than DL1r with bin_DL1r used in the training Similar with DL1r without bin_DL1r
1L	Similar for 50% and 60%	~4% lower than 70%, mainly in LP SR ~10% wores than DL1r with bin_DL1r used in the training Similar with DL1r without bin_DL1r
2L	Similar for all 3 WPs	Similar with DL1r with bin_DL1r in the training

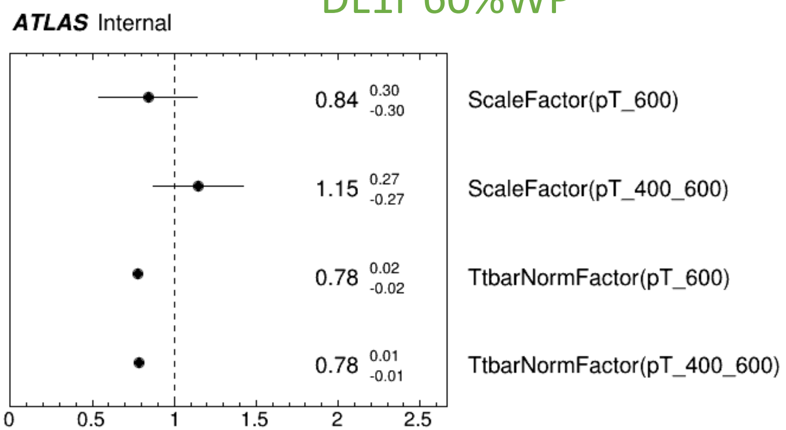
DL1r@70% VS. Xbb **[mBB]**

Channel	Xbb WPs	400_600GeV SR		600GeV SR	
		HP	LP	HP	LP
0-lepton channel	Xbb 60%	-2%	-26%	+150%	+25%
	Xbb 50%	+11%	-12%	+120%	+46%
1-lepton channel	Xbb 60%	+4%	+1%	-2%	+1%
	Xbb 50%	+8%	+8%	-3%	+9%
2-lepton channel	Xbb 60%	+11%		+24%	
	Xbb 50%	+9%		+26%	

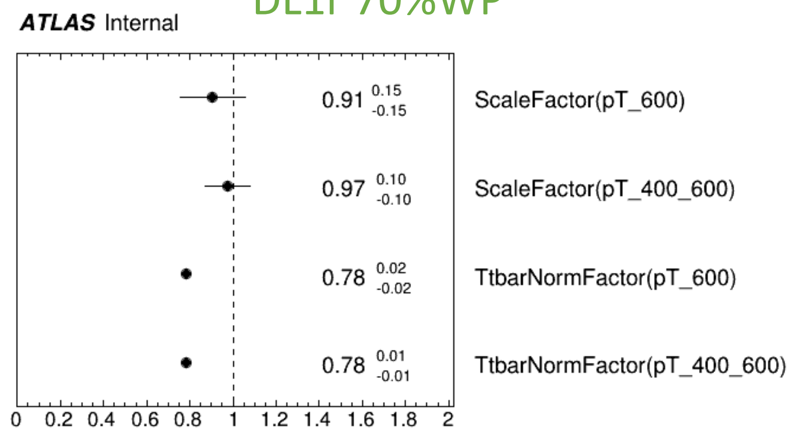
DL1r close-by effect study

Follow the Xbb ttbar calibration strategy:

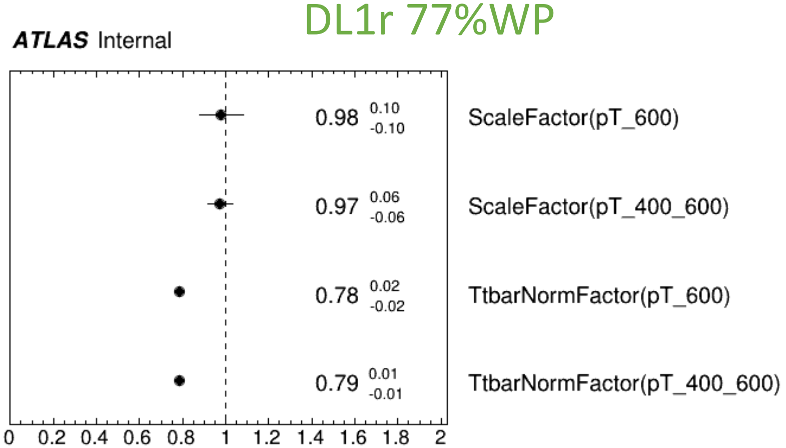
DL1r 60%WP



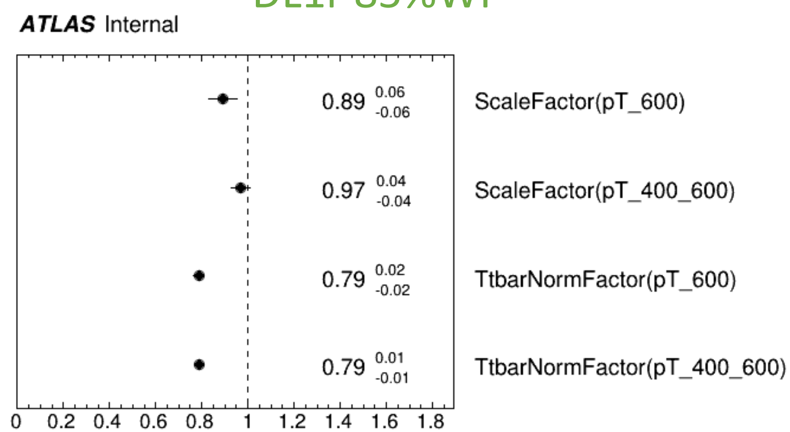
DL1r 70%WP



DL1r 77%WP

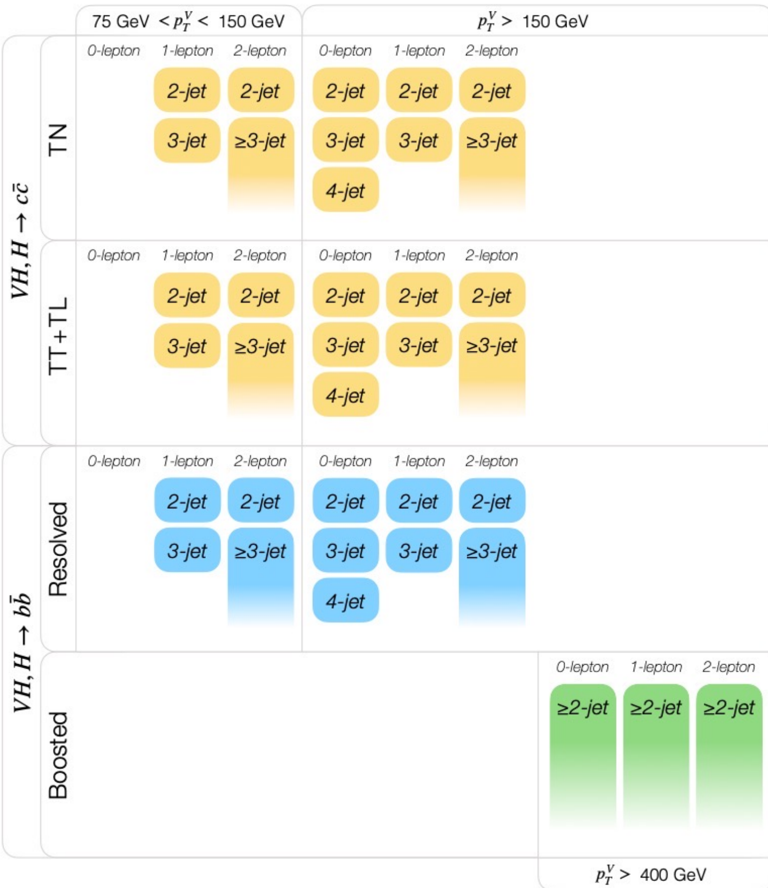


DL1r 85%WP



MVA overview

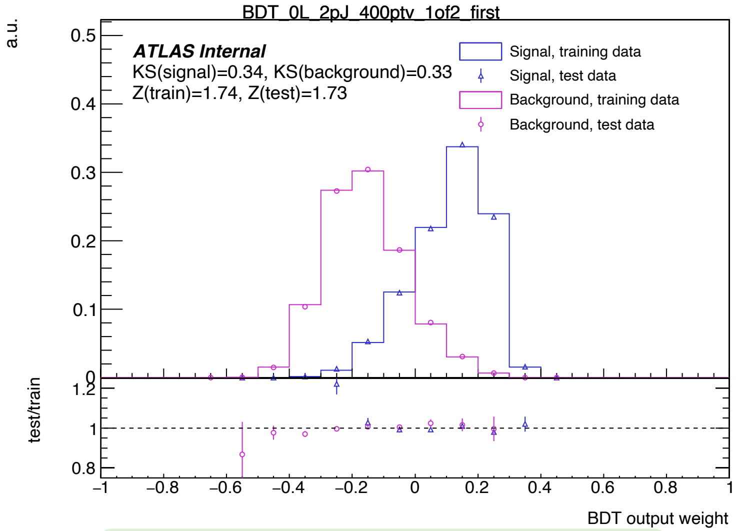
- The VH(bb/cc) analysis inherits the BDT method from the previous resolved VH(bb) and extends its usage to the boosted VH(bb) and VH(cc) regimes.
  - Boosted VH(bb): significance **+~50%**



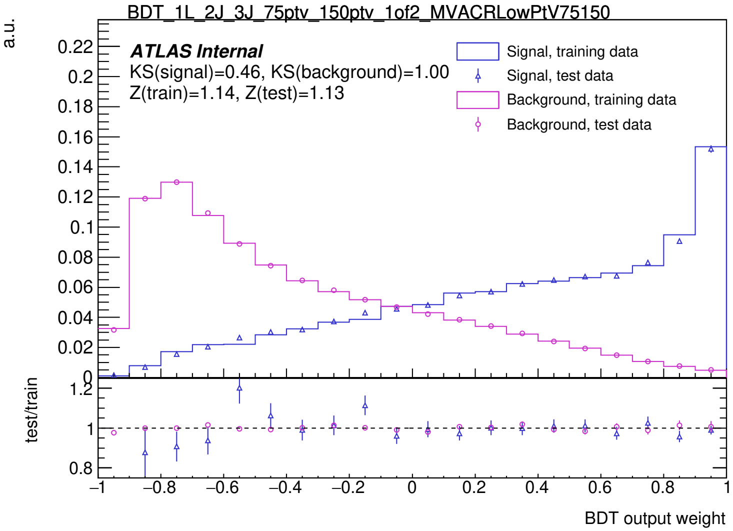
		$VH, H \to b\bar{b}/c\bar{c}$ Resolved			$VH, H \to b\bar{b}$ Boosted		
Variable	Name	0-lepton	1-lepton	2-lepton	0-lepton	1-lepton	2-lepton
$m_{j_1 j_2}$ or m_J	mBB / mJ	×	×	×	×	×	×
$m_{j_1 j_2 j_3}$	mBBJ	×	×	×			
$p_T^{j_1}$ or $p_T^{j_1,\mathrm{trk}}$	pTB1 / pTBTrkJ1	×	×	×	×	×	×
$p_T^{j_2}$ or $p_T^{j_2,\mathrm{trk}}$	pTB2 / pTBTrkJ2	×	×	×	×	×	×
$p_T^{j_3,\mathrm{trk}}$	pTBTrkJ3				×	×	×
$\sum_{i\neq 1,2} p_T^{j_i}$	sumPtAddJets	×	×	×			
$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ or $\Delta R(\boldsymbol{j_{1,\mathrm{trk}}}, \boldsymbol{j_{2,\mathrm{trk}}})$	dRBB / deltaRbTrkJbTrkJ	×	×	×	×	×	×
$\lvert\Delta\eta(\boldsymbol{j_1}, \boldsymbol{j_2})\rvert$	dEtaBB	×					
$\mathrm{bin}_{\mathrm{DL1r}}(j_1)$	bin_btagB1 / bin_bTagBTrkJ1	×	×	×	×	×	×
$\mathrm{bin}_{\mathrm{DL1r}}(j_2)$	bin_btagB2 / bin_bTagBTrkJ2	×	×	×	×	×	×
p_T^V	pTV	$\equiv E_T^{\mathrm{miss}}$	×	×	$\equiv E_T^{\mathrm{miss}}$	×	×
E_T^{miss}	MET	×	×		×	×	
$E_T^{\mathrm{miss}}/\sqrt{S_T}$	METSig			×			
$\lvert\Delta y(\boldsymbol{V}, \boldsymbol{H_{\mathrm{cand}}})\rvert$	dYVBB/deltaYVJ		×	×		×	×
$\lvert\Delta\phi(\boldsymbol{V}, \boldsymbol{H_{\mathrm{cand}}})\rvert$	dPhiVBB/absdeltaPhiVJ	×	×	×	×	×	×
$\min[\Delta\phi(\boldsymbol{\ell}, \boldsymbol{b}$ or $\boldsymbol{c})]$	dPhiLBmin		×				
m_{eff}	MEff	×					
m_T^W	mTW		×				
m_{top}	Mtop		×				
$m_{\ell\ell}$	mLL			×			
$\cos\theta(\boldsymbol{\ell^-}, \boldsymbol{Z})$	cosThetaLep			×			×
p_T^{ℓ}	pTL					×	
N(track-jets in J)	NAdditionalCaloJets				×	×	×
N(add. small R-jets)	NMatchedTrackJetLeadFatJet				×	×	×
Colour	Colour				×	×	×
$\min\{\Delta R(b, j)\}$	minDRBjets	×	×				

MVA: Overtraining check

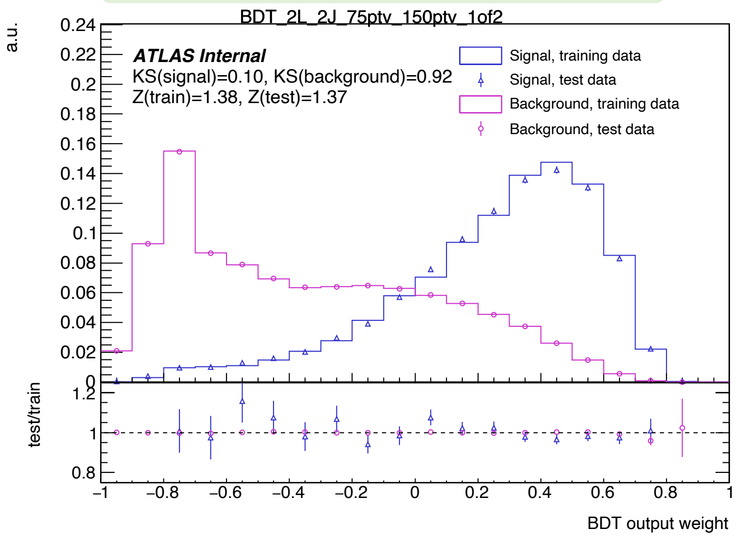
VH training (0L)



mvaCRLow training (1L)



VZ training (2L)



No obvious overtraining issue for all trainings.

MVA: hyperparameters

- The hyperparameters of the BDTs have been optimized per region of for each leptonic channel to **reduce overtraining** and **maximize the BDT statistic only significance.**

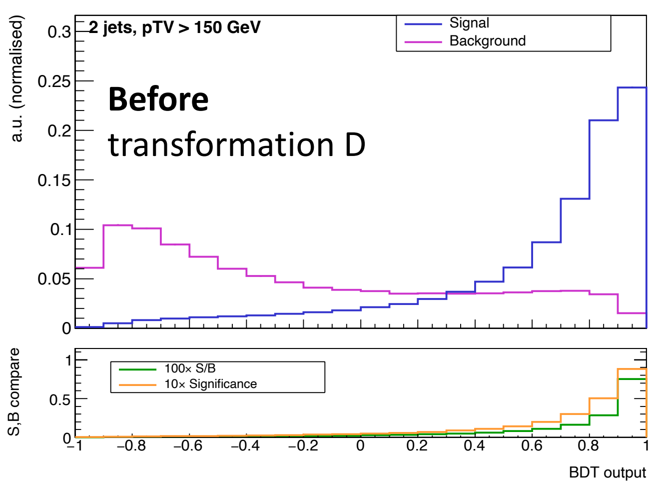
	$VH, H \to c\bar{c}$		$VH, H \to c\bar{c}$, Diboson as signal
Settings	0-, 1- and most 2-lepton regions	2-leton, 3pJ, low pTV	0-, 1- and 2-lepton
Boost type	Gradient boost	Adaboost	Adaboost
Number of trees	600	200	200
Maximum depth	4	4	4
Learning rate (β)	0.5	0.15	0.15
Number of cuts	100	100	100
Minimum node size	5%	5%	5%
Separation method	Gini index	Gini index	Gini index
Pruning method	No pruning	No pruning	No pruning

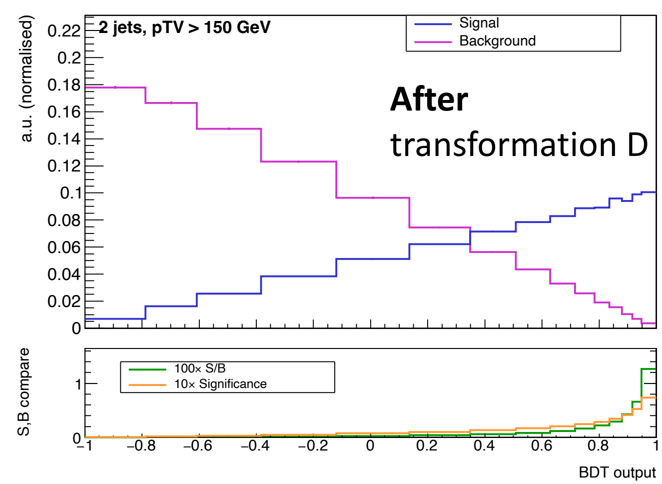
	Resolved $VH, H \to b\bar{b}$			Boosted $VH, H \to b\bar{b}$		
Settings	0-lepton	1-lepton	2-lepton	0-lepton	1-lepton	2-lepton
Boost type	Gradient boost	Gradient boost	Gradient boost	Adaboost	Adaboost	Adaboost
Number of trees	200	600	200	800	800	400
Maximum depth	3	4	4	3	3	3
Learning rate (β)	0.5	0.5	0.5	0.35	0.35	0.3
Number of cuts	100	100	100	60	60	100
Minimum node size	5%	5%	5%	2%	2%	7%
Separation method	Gini index	Gini index	Gini index	Gini index	Gini index	Gini index
Pruning method	No pruning	No pruning	No pruning	No pruning	No pruning	No pruning

MVA: output transformation

- The originlal BDT distribution have 500 equidistant bins. A transformation D* is adpoted for BDT rebining to improve sensitivity.
- The parameters of transformation D are optimized to ensure enough statistics in each bin and at the same time ensuring sensitivity to signal.

	$p_T^V \in [75, 150]$ GeV	$p_T^V \in [150, 250]$ GeV	$p_T^V \in [250, 400]$ GeV	$p_T^V \in [400, 600]$ GeV	$p_T^V > 600$ GeV
$VH, H \to b\bar{b}$	$z_s = 10,\ z_b = 5$		$z_s = 5,\ z_b = 3$		$z_s = \begin{cases} 3 & \text{for 0-, 1-lepton} \\ 2 & \text{for 2-lepton} \end{cases},\ z_b = 2$
$VH, H \to c\bar{c}$	$\begin{cases} \text{TT-tagged: } z_s = 5, z_b = 3 \\ \text{Else: } z_s = 10,\ z_b = 5 \end{cases}$	$\begin{cases} \text{0-/1-lepton} & \begin{cases} \text{TT-tagged: } z_s = 5, z_b = 3 \\ \text{Else: } z_s = 10,\ z_b = 5 \end{cases} \\ \text{2-lepton} & \begin{cases} \text{TT-tagged: } z_s = 2,\ z_b = 2 \\ \text{LT-/XT-tagged: } z_s = 5,\ z_b = 5 \\ \text{Else: } z_s = 10,\ z_b = 5 \end{cases} \end{cases}$	$\begin{cases} \text{TT-tagged: } z_s = 2, z_b = 2 \\ \text{LT-/XT-tagged: } z_s = 5,\ z_b = 3 \\ \text{Else: } z_s = 10,\ z_b = 5 \end{cases}$		

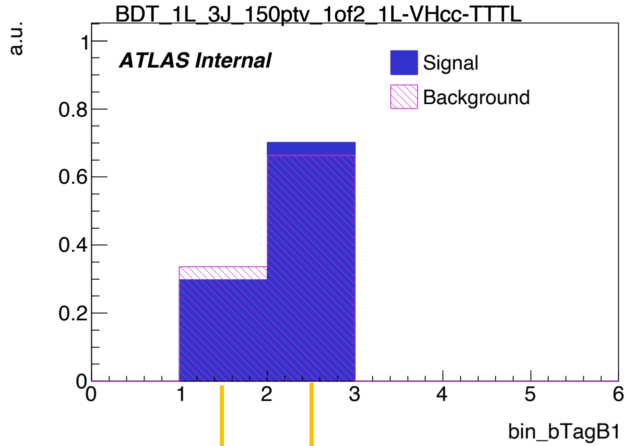




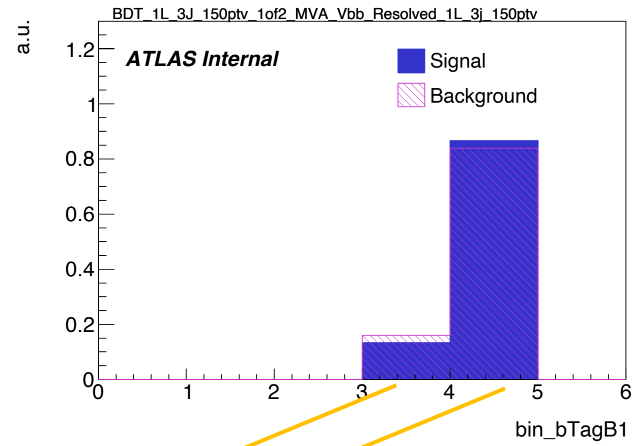
*Formula used for transformation D is: $Z = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b}$.

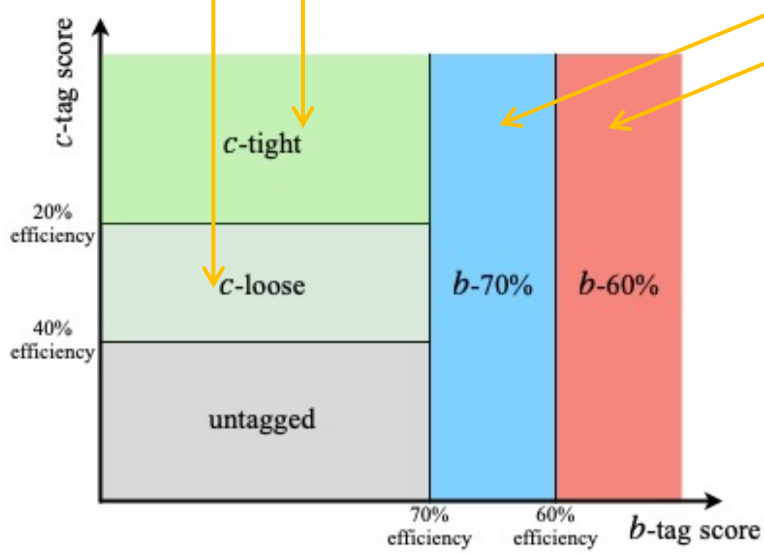
MVA: new variables

VH(cc)

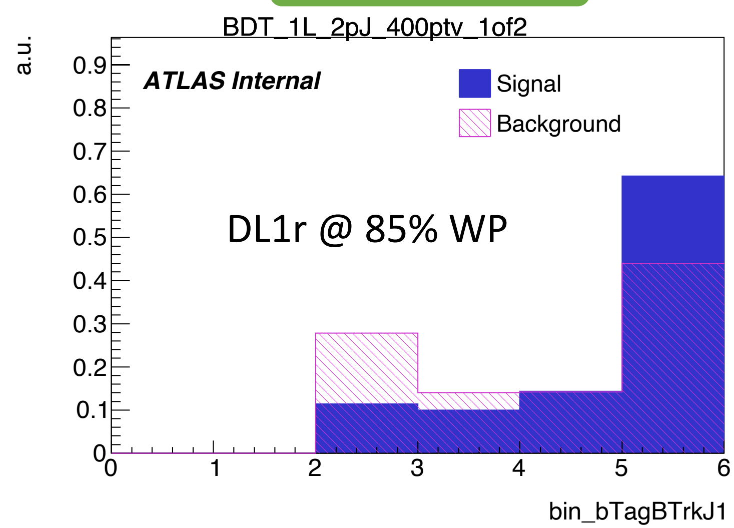


Resolved VH(bb)

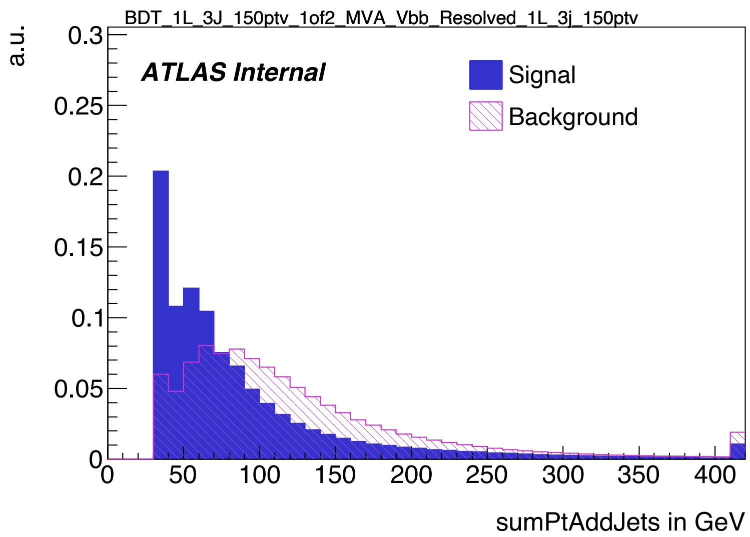


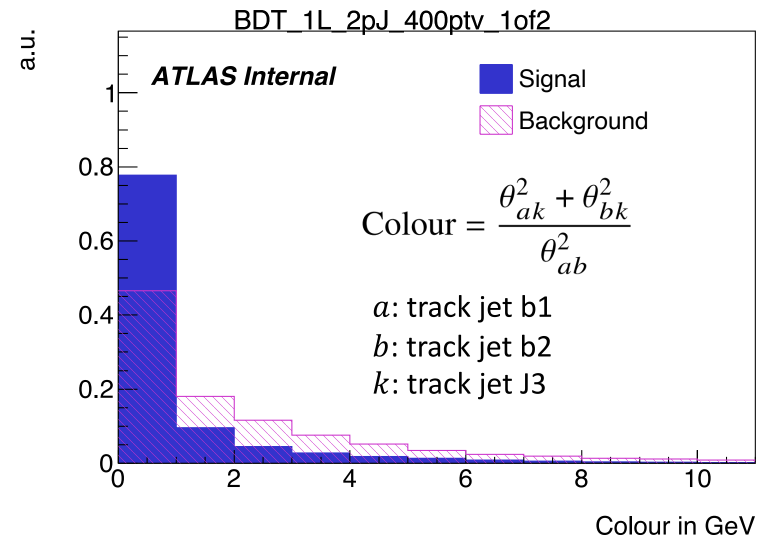


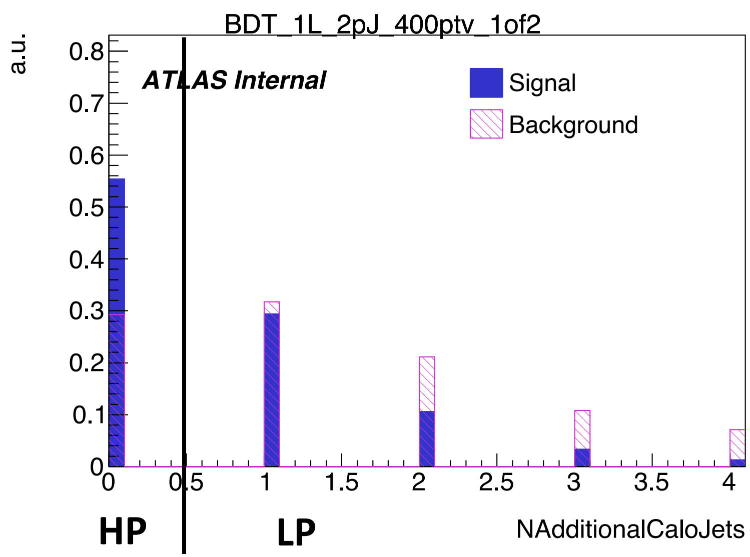
Boosted VH(bb)

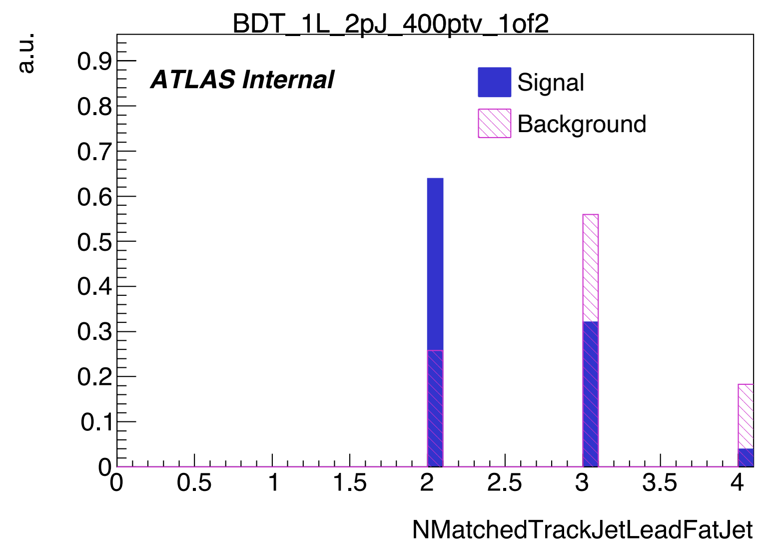


MVA: new variables



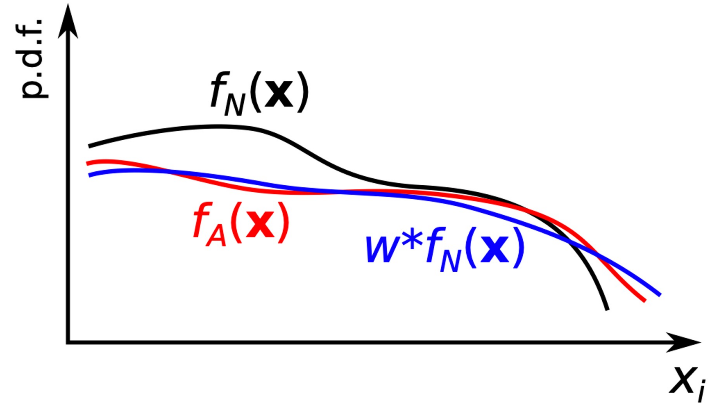


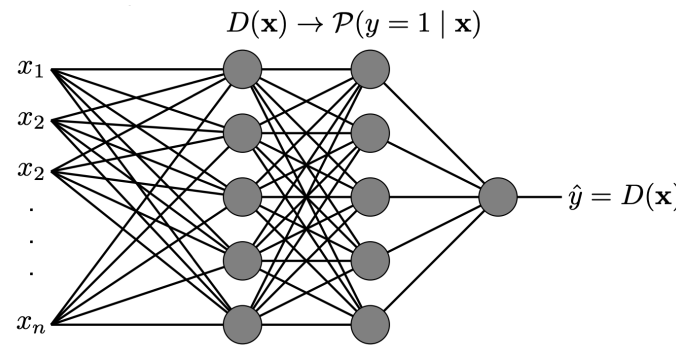


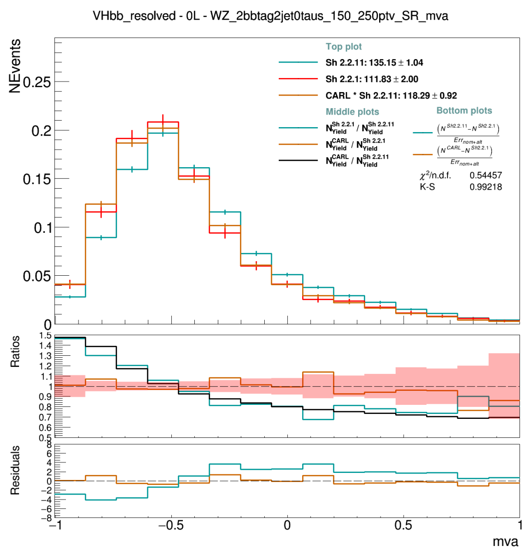


CARL re-weighting technique

- **Re-weighting**: take events sampled from the nominal and re-weight them event-by-event to make the re-weighted distribution looks like the alternative.
  - Motivated by low statistics of alternative sample
- **CARL** is a re-weighting technique based on **DNN**
  - Simultaneously reweight multiple variables, and **correlations** between them are considered
  - Important as now SRs fitted with BDT
  - Replace the previous BDTr technique
- The output of the neural network D(x) approaches probability that event is alternative







Top panel:
- Green: Nominal (Sh2.2.11) sample
- Red: Alternative (Sh2.2.1) sample
- Orange: reweighted nominal

Middle panel: ratios
- Green (Black): ratio of alternative (Nominal * CARL) over the nominal
- Orange: ratio of Nominal * CARL over the alternative sample --> **should be at 1**

Bottom panel: residual = (ratio - 1)/relative error
- Green: residual of nominal
- Orange: residual of Nominal*CARL --> **should be at 0**

Breakdown

```
Set of nuisance          Impact on error
parameters
----------------------------------------------------------
  Total                       + 0.133  - 0.129  +- 0.131
 DataStat                     + 0.097  - 0.096  +- 0.097
  FullSyst                    + 0.091  - 0.087  +- 0.089
 Data stat only               + 0.088  - 0.087  +- 0.088
 Top-emu CR stat              + 0.011  - 0.010  +- 0.010
 Floating normalizations      + 0.031  - 0.031  +- 0.031
 Modelling: VH                + 0.024  - 0.019  +- 0.021
 Modelling: Background        + 0.070  - 0.069  +- 0.069
 Multi Jet                    + 0.004  - 0.004  +- 0.004
 Modelling: single top        + 0.016  - 0.015  +- 0.015
 Modelling: ttbar             + 0.019  - 0.019  +- 0.019
 Modelling: W+jets            + 0.031  - 0.031  +- 0.031
 Modelling: Z+jets            + 0.035  - 0.035  +- 0.035
 Modelling: Diboson           + 0.026  - 0.025  +- 0.025
 MC stat                      + 0.018  - 0.018  +- 0.018
 Experimental Syst            + 0.044  - 0.040  +- 0.042
 Detector: lepton             + 0.005  - 0.005  +- 0.005
 Detector: MET                + 0.006  - 0.005  +- 0.005
 Detector: JET                + 0.034  - 0.030  +- 0.032
Detector: FATJET                + 0.012  - 0.011  +- 0.012
Detector: FTAG PFlow (b-jet)       + 0.018  - 0.017  +- 0.017
Detector: FTAG PFlow (c-jet)       + 0.011  - 0.011  +- 0.011
Detector: FTAG PFlow (l-jet)       + 0.003  - 0.003  +- 0.003
Detector: FTAG PFlow (extrap)      + 0.000  - 0.000  +- 0.000
Detector: FTAG VR (b-jet)       + 0.005  - 0.005  +- 0.005
 Detector: PU                 + 0.007  - 0.006  +- 0.007
 Lumi                         + 0.008  - 0.006  +- 0.007
```