Searching and measuring the beauty and the charm of the Higgs boson with ATLAS detector

Yanhui Ma

Higgs Potential 2024 @Hefei
2024.12.20

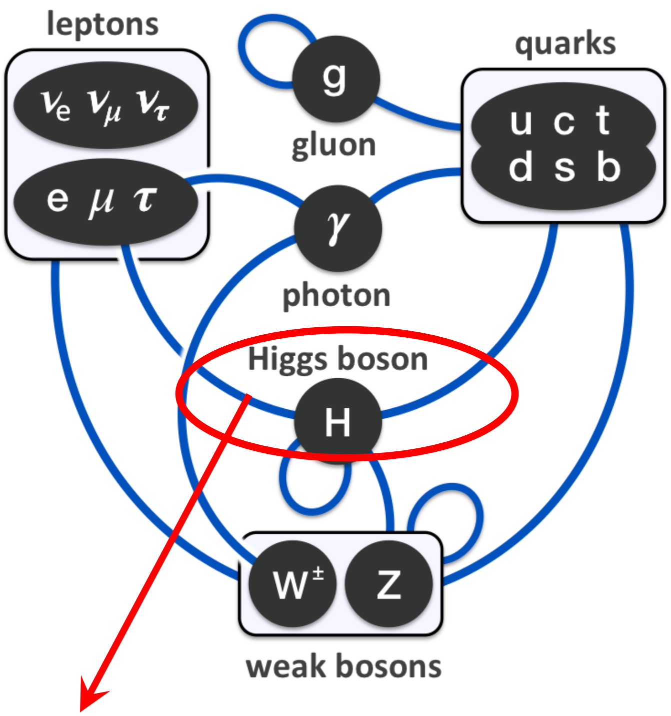
ATLAS

OKLAHOMA STATE UNIVERSITY
1890

The Higgs boson in the Standard Model (SM)

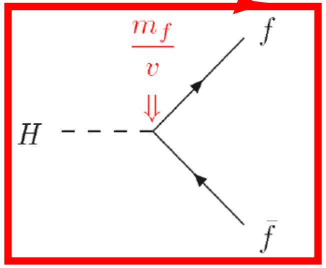
- The SM is the most thoroughly tested theory of particle physics that has had a great success to explain experimental observations of particle physics.

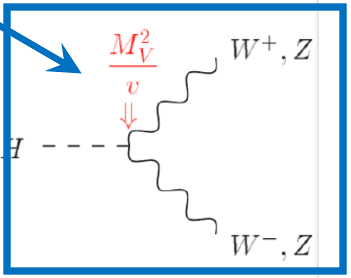
- In the SM, the Higgs mechanism provides masses to bosons and fermions



$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + h.c. + \bar{\psi}_i y_{ij}\psi_j\phi + h.c. + |D_\mu\phi|^2 - V(\phi)$$

Discovered in 2012 with a mass ~125 GeV

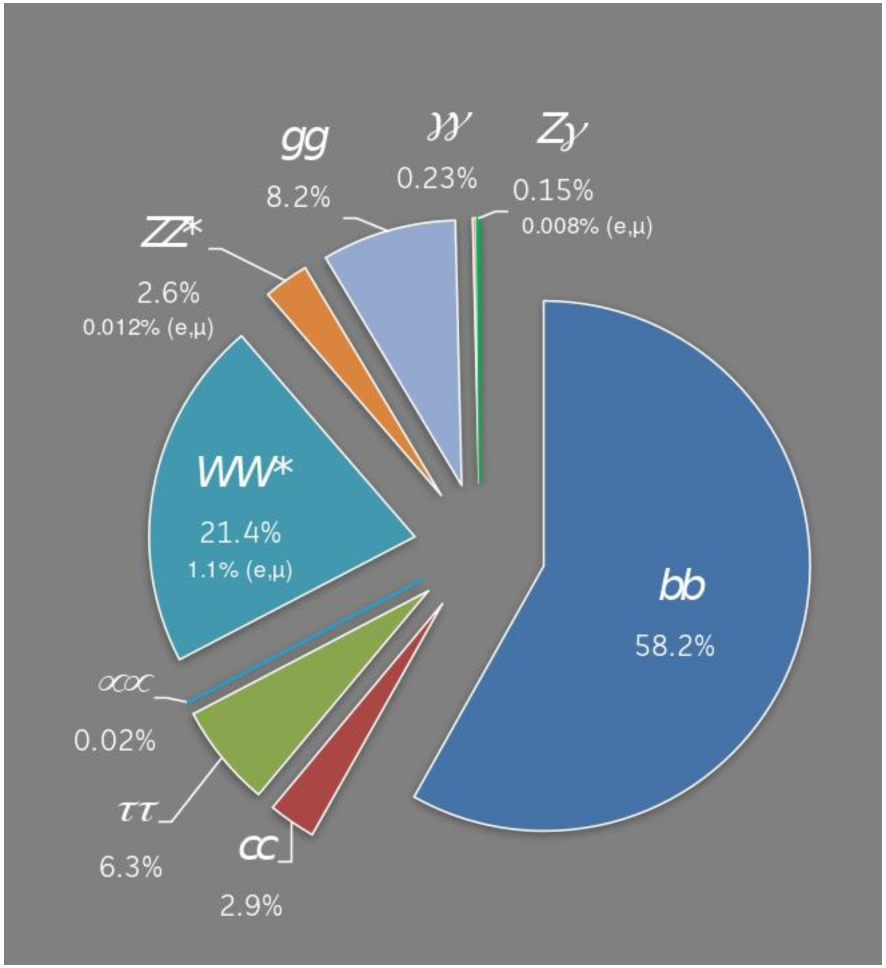




Higgs boson decays

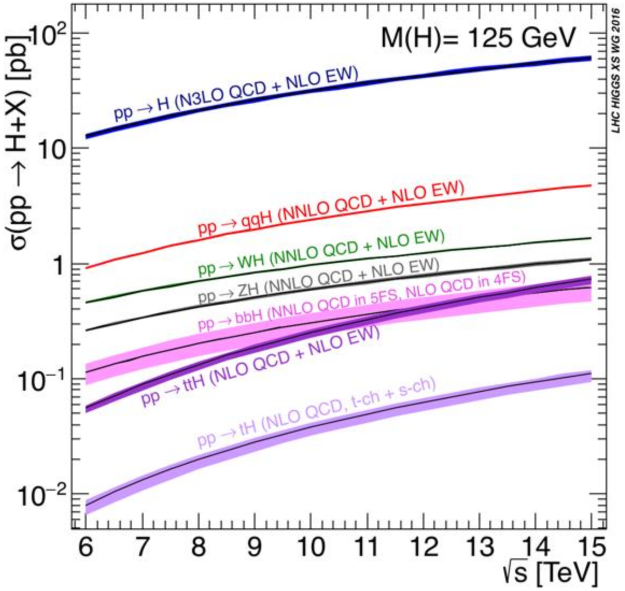
- Many decay modes accessible at the LHC
- bb largest BR ~ 58%
  - Measurement of the Yukawa coupling to down type quarks.
  - Constrain the Higgs boson decay width.
  - If BSM particles allowed in loops and decays: Measuring H→bb limits BSM branching fraction allowed
- cc BR ~3%
  - Probe of Higgs coupling to 2nd generation of quarks
  - One of the largest contribution to Higgs width that we have no evidence for
  - Small charm Yukawa coupling → susceptible to significant modifications in various new physics scenarios

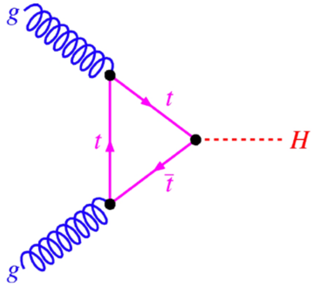
Higgs boson branching ratios



Higgs boson production at hadron colliders

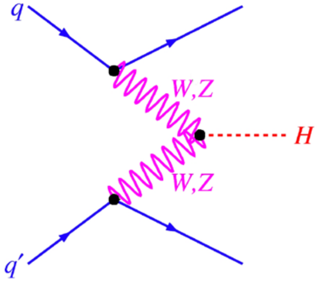
- 4 main channels at the LHC
- Total cross section 56pb at 13 TeV
  - ~6 million Higgs bosons produced in ATLAS





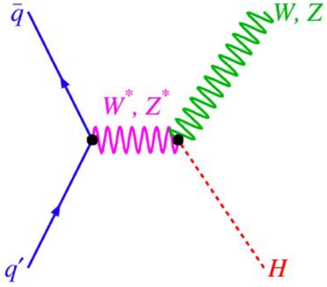
Gluon fusion (ggF)

- Dominant mode (88% of the total)



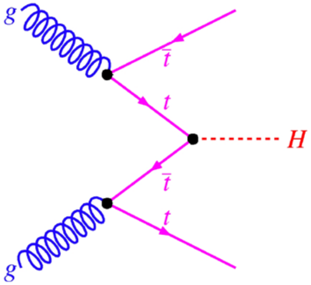
Vector boson fusion(VBF)

- 7% of the total



VH (WH/ZH)

- 3% of the total



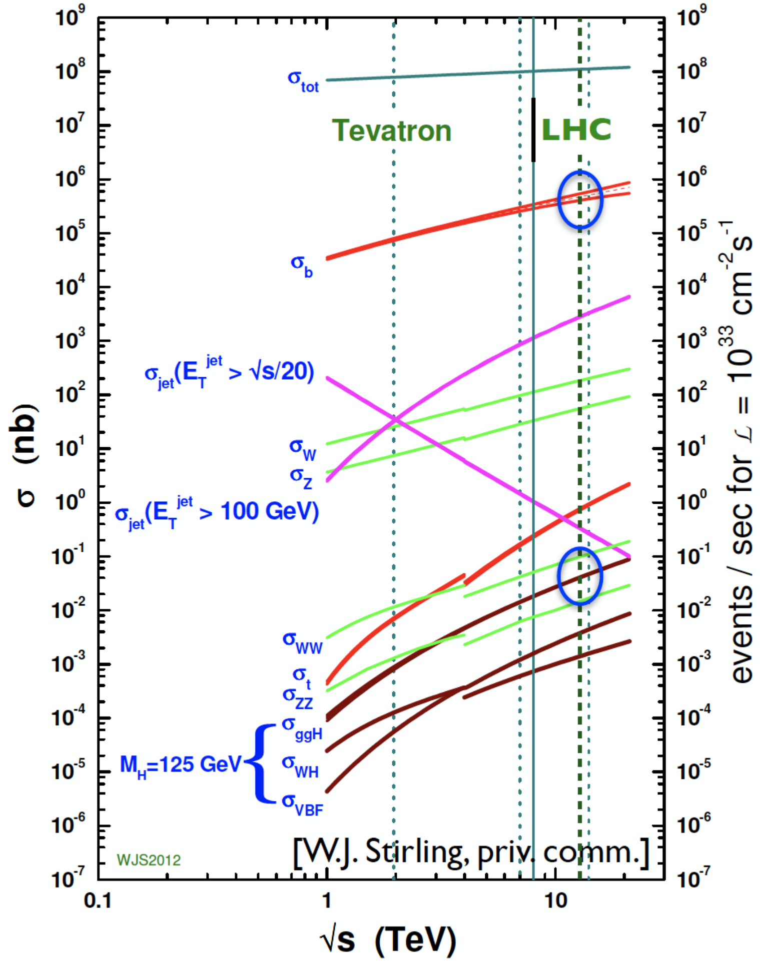
ttH

- 1% of the total

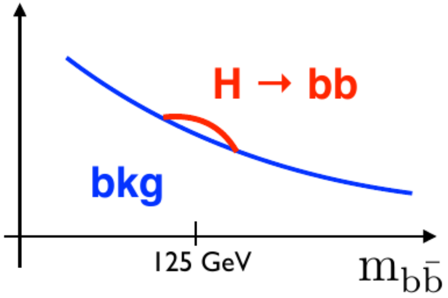
VH: golden channel to study the Hbb and Hcc decay

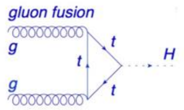
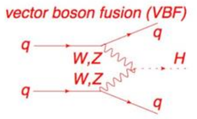
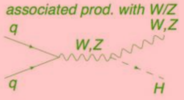
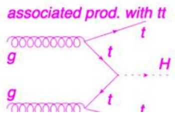
- Very large multi-b/c-jets production cross section at the LHC

proton - (anti)proton cross sections



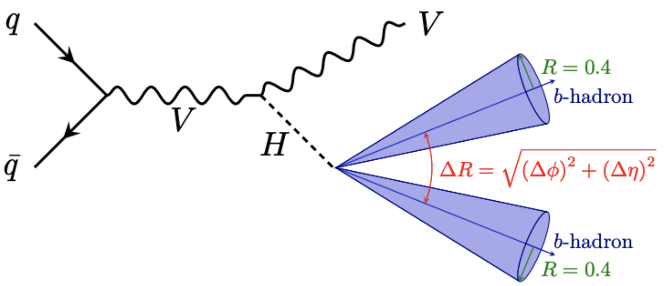
- Inclusive Higgs boson production (2 b/c-jets in final state) overwhelmed by bkgs by many orders of magnitude

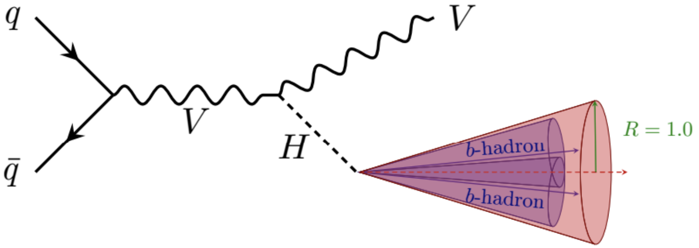


Production mode	Primary signature	
	Just H → bb	• Huge multi-jet background • Triggering possible at high p_T(H), but S/B expected to be ~ O(0.1%) • Jet substructure analysis by CMS (p_T(H)>450 GeV)
	2 VBF jets (+ γ)	• Large multi-jet background • Still a fully hadronic final state: trigger and background modeling is challenging • Additional γ helps (~similar sensitivity, higher S/B)
	W, Z	• Exploit leptonic signatures for trigger, and suppression of multi-jet background. • *Main search channel for H → bb at the LHC!*
	top+anti-top	• Leptonic signatures for trigger, but challenging due to combinatorics and tt+bb backgrounds • But gives access also to top quark coupling!

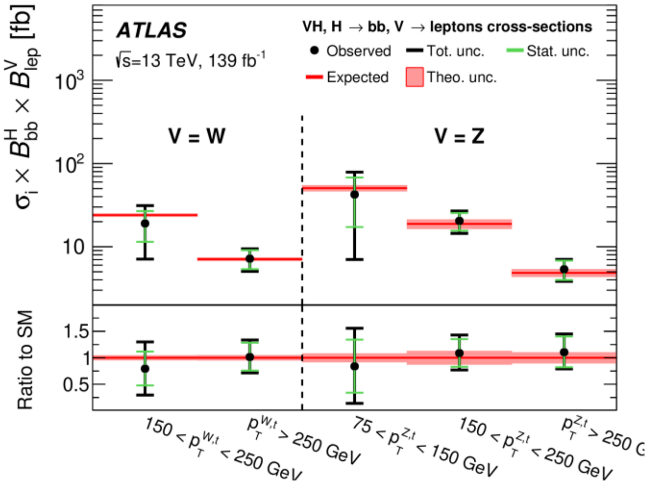
Where did ATLAS stand in 2022

- Three standalone VHbb/cc results with ***full Run 2 dataset***:
  - Resolved VHbb
    - Reconstructing the Higgs candidate with two small R (resolved) jet
    - Observation of VHbb well established with more than > 5 sigma, focused more on the cross-section measurement
  - Boosted VHbb
    - reconstructing the Higgs candidate with one large R (boosted) jet
    - First attempt for studying the Hbb decay in the extreme high pTV regime (pTV > 250 GeV)
  - VHcc
    - Considered only the resolved regime

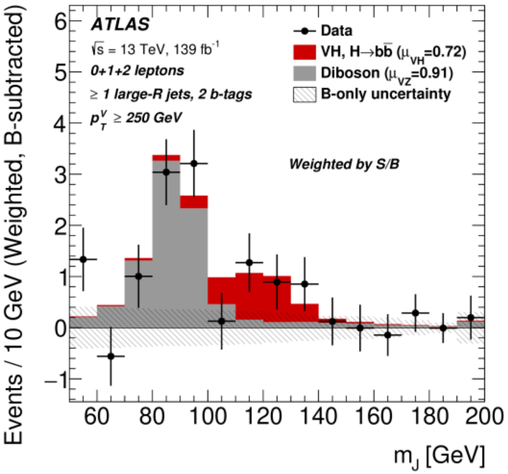




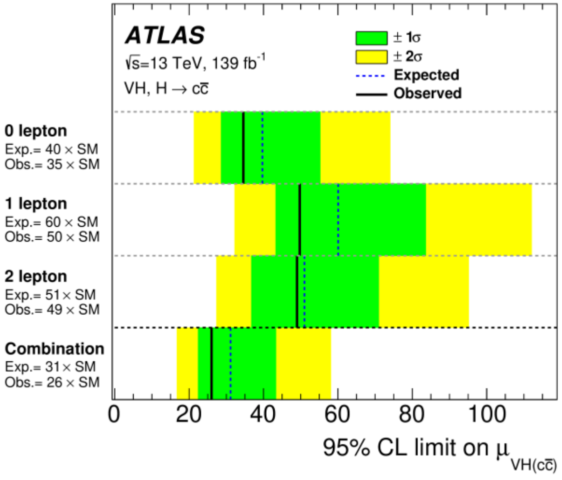
Resolved VHbb



Boosted VHbb



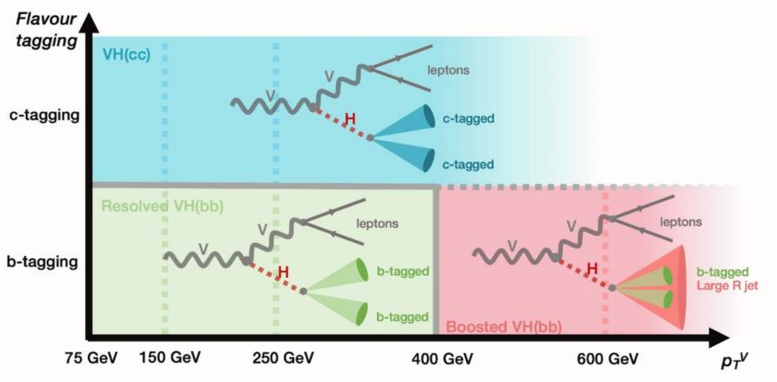
VHcc



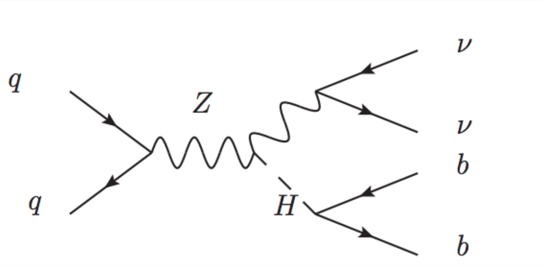
The legacy VHbb/cc analysis

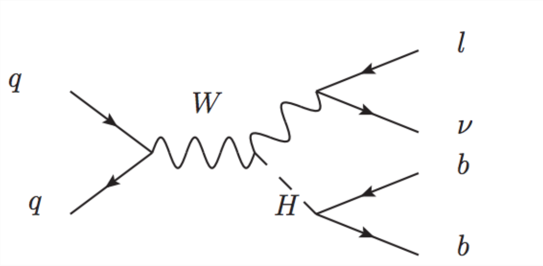
arXiv:2410.19611

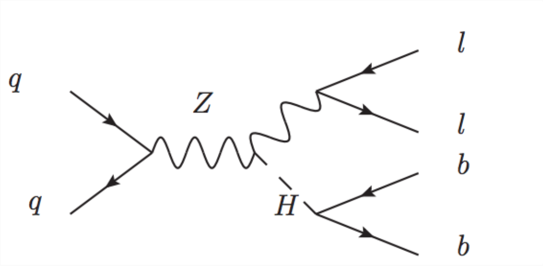
- A combined measurement for resolved VH(bb), boosted VH(bb) and VH(cc) analysis
  - Takig advantage of the similarities in the experimental signatures of H→bb and H→cc decays
  - VHbb resolved and boosted regiem separation at pTV = 400 GeV
  - VHcc consider only the reloved regiem



- Leptonic decays (electron or muon) of W/Z for background rejection and trigger → 3 channels : 0,1,2 leptons.



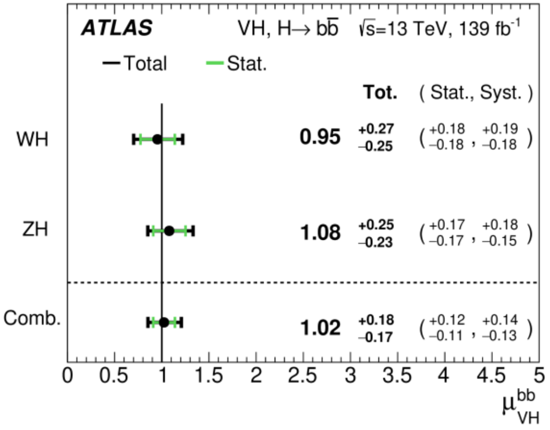


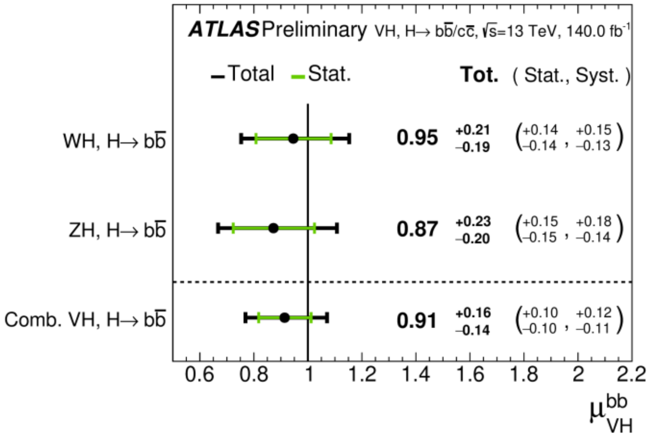


At a glance --- The much-improved results

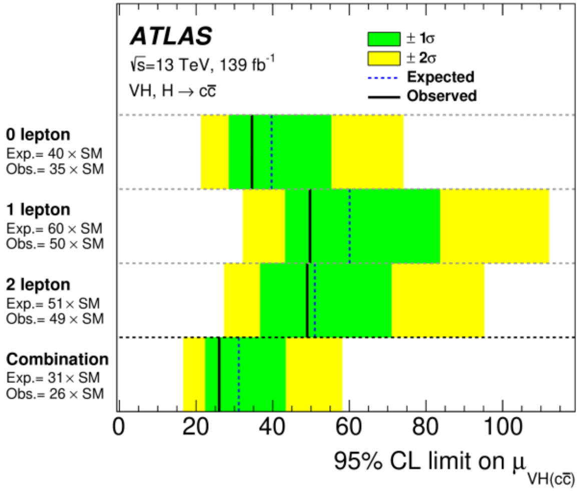
First full run 2 results

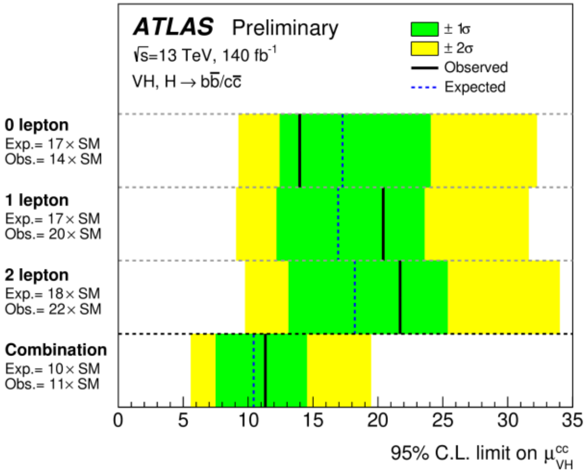
The legacy analysis results





- VHbb: ~10-20% improvement on the VH cross section measurements. Most precise measurement to date

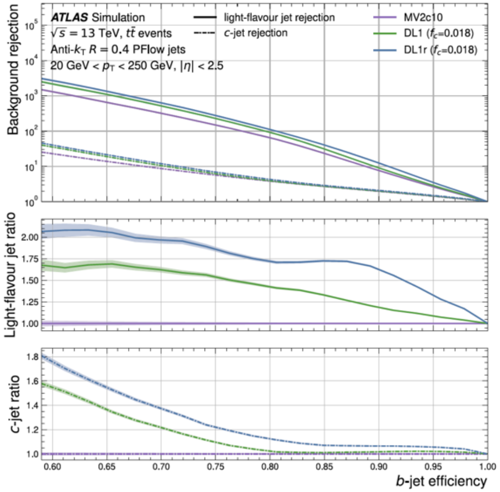


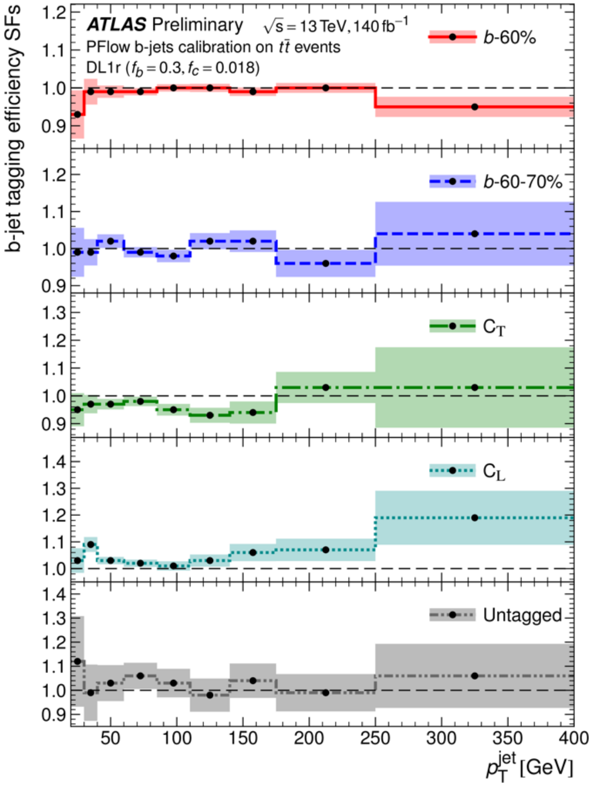


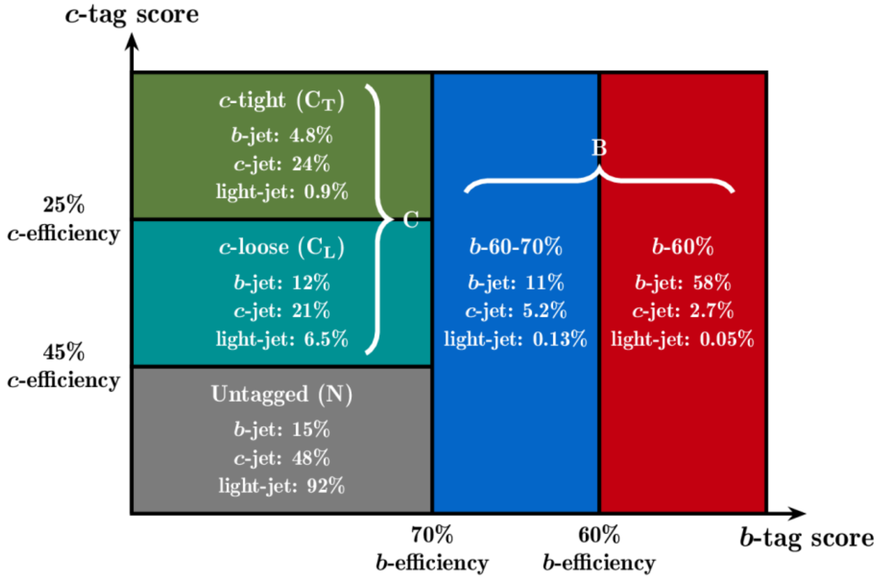
- VHcc: ~ a factor of 3 improvement on the upper limit setting. Strongest observed limit to date

The key ingredients for the improvements – flavor-tagging

- Moved from the BDT based MV2 tag to the deep neural network based DL1r tag, with a lot of improved/ newly introduced algorithms
  - more than 20% better c/light-jet rejection
- Dedicated optimization for the definition of the B/C-tag working point
  - making sure the orthogonality of the b and c jet regions
- Dedicated calibrations performed in these WP.
  - precision ~10% for c-jet and ~3% for b-jets

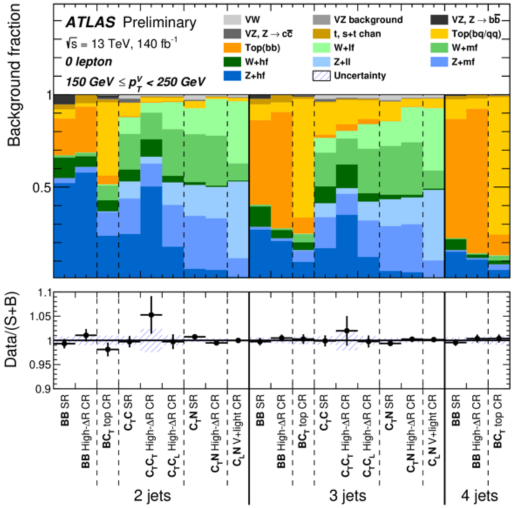


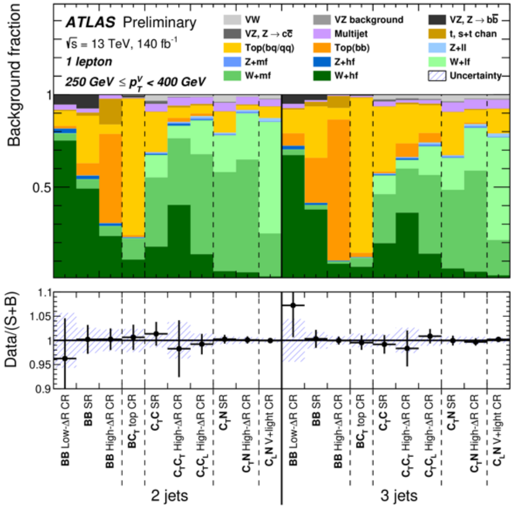


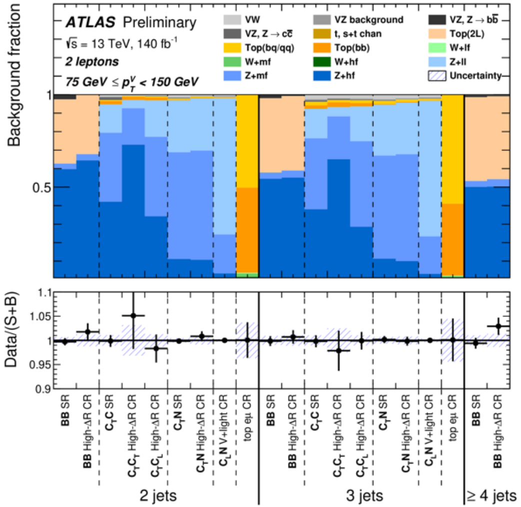


The key ingredients for the improvements – background estimation

- Achieve a good control of the very diverse backgrounds across the different lepton channels and tag regions is the key challenge of the analysis
- Main background: V+jets
  - ***Z+jets*** dominant in 0- and 2-lepton
  - ***W+jets*** dominate in 0 and 1-lepton
- Main background: Top quark processes (ttbar + single top Wt)
  - Top (***bb*** and ***bqqq***) dominant in 0- and 1-lepton and VHbb 2-lepton
- Other small backgrounds: ***single top t and s channel***; ***diboson***; ***multijet*** (1L)

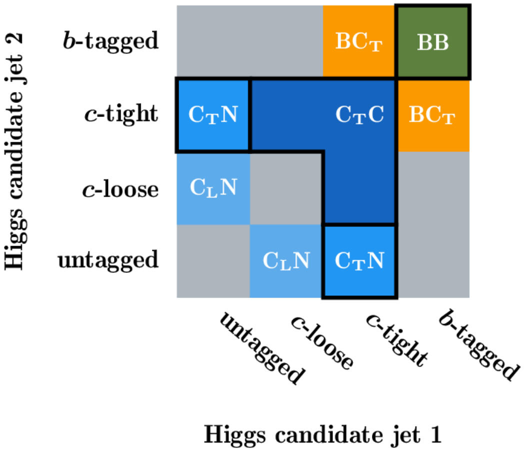


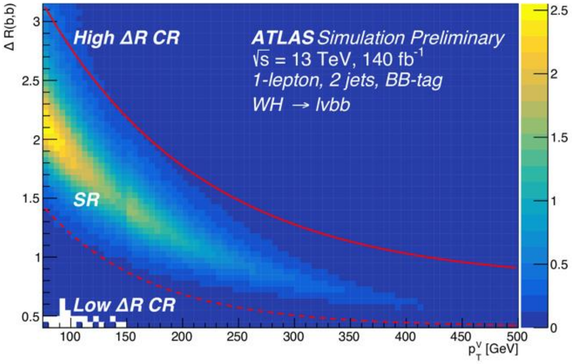


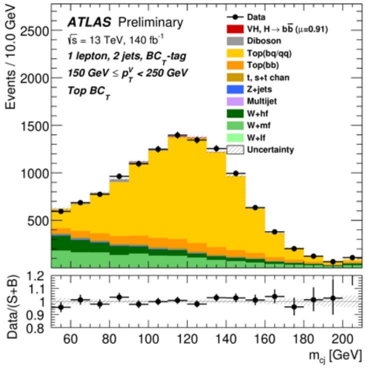


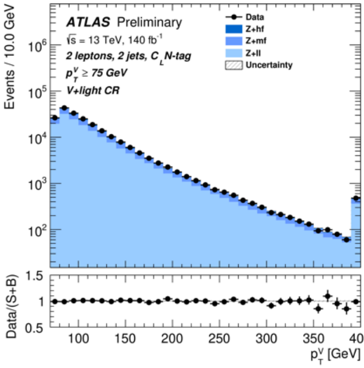
The key ingredients for the improvements – background estimation

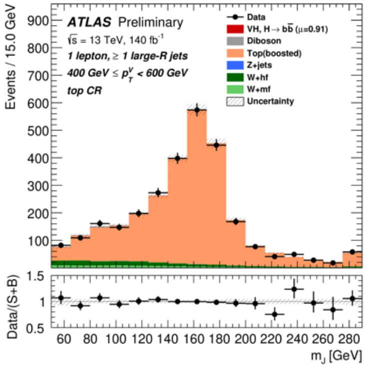
- More than 50 SRs considered:
  - VHbb: request two b-tagged jets (BB tagged)
  - VHcc: two c-tagged jets (C_TC) and 1 c-tagged jet (C_TN)
  - Split in different pTV and nJets regiems
- More than 100 dedicated high purity CRs defined to better constant the normalization and shape of the main backgrounds
  - BC_T -> Top (bc) CR
  - C_LN -> V+light jet CR
  - Use jet-angular separation ΔR(b,b) defining CR enriched in: V+bb (Low dR CR); Top(bb) (high dR CR)





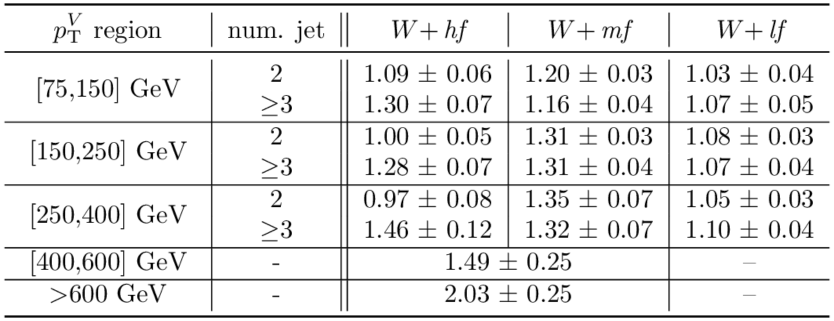




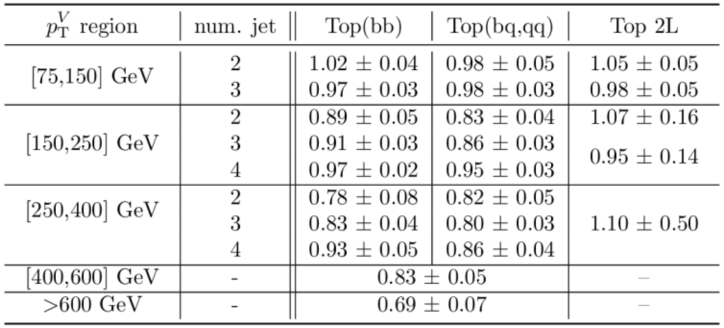


The key ingredients for the improvements – background estimation

- Main bkgs normalization are well constrained by the CRs and are floating in the fit
  - The tends of the NFs are in good agreements with the dedicated measurements (V+jets; ttbar)



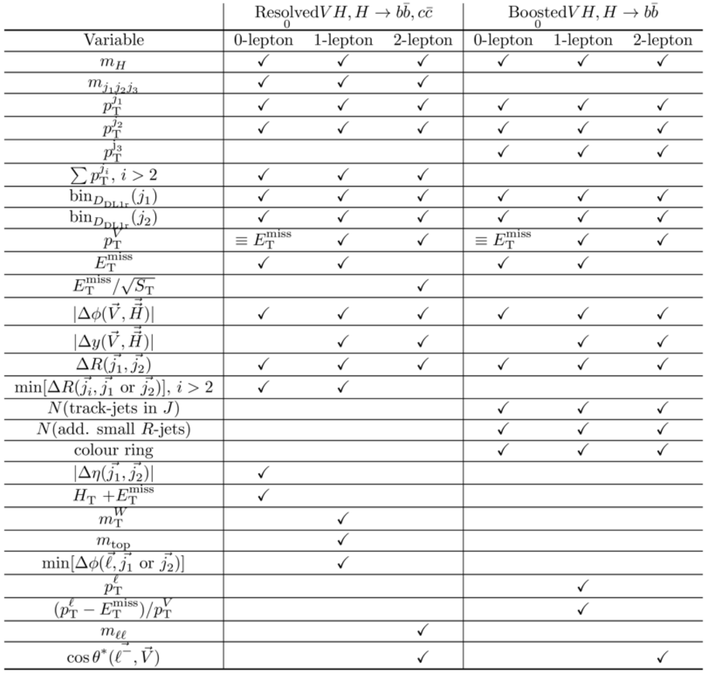
p_T^V region	num. jet	$W+hf$	$W+mf$	$W+lf$
[75,150] GeV	2	1.09 ± 0.06	1.20 ± 0.03	1.03 ± 0.04
	≥3	1.30 ± 0.07	1.16 ± 0.04	1.07 ± 0.05
[150,250] GeV	2	1.00 ± 0.05	1.31 ± 0.03	1.08 ± 0.03
	≥3	1.28 ± 0.07	1.31 ± 0.04	1.07 ± 0.04
[250,400] GeV	2	0.97 ± 0.08	1.35 ± 0.07	1.05 ± 0.03
	≥3	1.46 ± 0.12	1.32 ± 0.07	1.10 ± 0.04
[400,600] GeV	-	1.49 ± 0.25		–
>600 GeV	-	2.03 ± 0.25		–



p_T^V region	num. jet	Top(bb)	Top(bq,qq)	Top 2L
[75,150] GeV	2	1.02 ± 0.04	0.98 ± 0.05	1.05 ± 0.05
	3	0.97 ± 0.03	0.98 ± 0.03	0.98 ± 0.05
[150,250] GeV	2	0.89 ± 0.05	0.83 ± 0.04	1.07 ± 0.16
	3	0.91 ± 0.03	0.86 ± 0.03	0.95 ± 0.14
	4	0.97 ± 0.02	0.95 ± 0.03	
[250,400] GeV	2	0.78 ± 0.08	0.82 ± 0.05	1.10 ± 0.50
	3	0.83 ± 0.04	0.80 ± 0.03	
	4	0.93 ± 0.05	0.86 ± 0.04	
[400,600] GeV	-	0.83 ± 0.05		–
>600 GeV	-	0.69 ± 0.07		–

The key elements for the improvements – ML techniques

- Improving the signal and background separation
  - simple and robust Boosted Decision Tree (BDT) is used achieve the maximum separation of signal and background
  - Input variables and training parameters tuned to yield best sensitivity, various checks performed to make sure the training works well.
  - For the first time BDT also applied to VHcc and boosted VHbb, yields >80% sensitivity improvements
  - The other “more fancy” methods (DNN, etc) also tested and resulted very similar performance with BDT

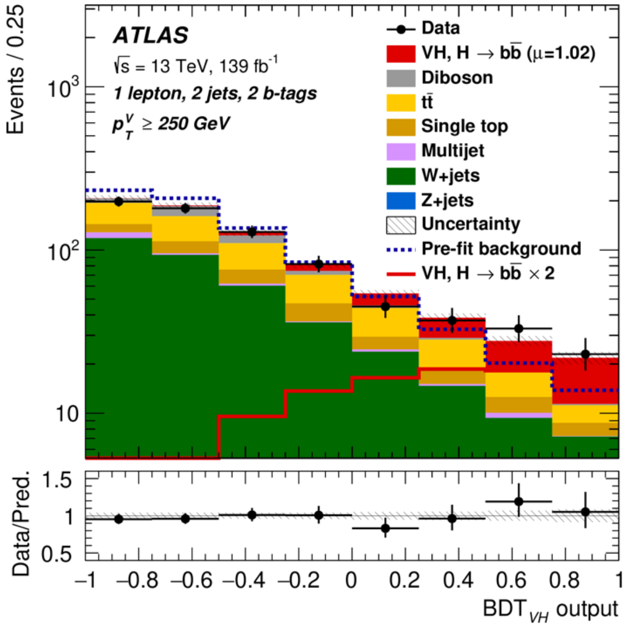


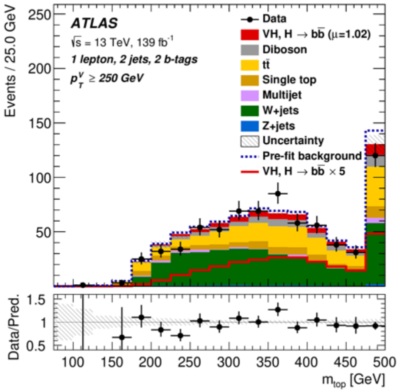
	Res$\underset{0}{\mathrm{o}}$lved$VH, H \to b\bar{b}, c\bar{c}$			Boo$\underset{0}{\mathrm{s}}$ted$VH, H \to b\bar{b}$		
Variable	0-lepton	1-lepton	2-lepton	0-lepton	1-lepton	2-lepton
m_H	✓	✓	✓	✓	✓	✓
$m_{j_1 j_2 j_3}$	✓	✓	✓			
$p_{\mathrm{T}}^{j_1}$	✓	✓	✓	✓	✓	✓
$p_{\mathrm{T}}^{j_2}$	✓	✓	✓	✓	✓	✓
$p_{\mathrm{T}}^{\mathrm{j}_3}$				✓	✓	✓
$\sum p_{\mathrm{T}}^{j_i}$, $i > 2$	✓	✓	✓			
$\mathrm{bin}_{D_{\mathrm{DL1r}}}(j_1)$	✓	✓	✓	✓	✓	✓
$\mathrm{bin}_{D_{\mathrm{DL1r}}}(j_2)$	✓	✓	✓	✓	✓	✓
p_{T}^{V}	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓
$E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓		✓	✓	
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$			✓			
$\lvert\Delta\phi(\vec{V}, \vec{H})\rvert$	✓	✓	✓	✓	✓	✓
$\lvert\Delta y(\vec{V}, \vec{H})\rvert$		✓	✓		✓	✓
$\Delta R(\vec{j_1}, \vec{j_2})$	✓	✓	✓	✓	✓	✓
$\min[\Delta R(\vec{j_i}, \vec{j_1}$ or $\vec{j_2})]$, $i > 2$	✓	✓				
N(track-jets in J)				✓	✓	✓
N(add. small R-jets)				✓	✓	✓
colour ring				✓	✓	✓
$\lvert\Delta\eta(\vec{j_1}, \vec{j_2})\rvert$	✓					
$H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$	✓					
m_{T}^{W}		✓				
m_{top}		✓				
$\min[\Delta\phi(\vec{\ell}, \vec{j_1}$ or $\vec{j_2})]$		✓				
p_{T}^{ℓ}					✓	
$(p_{\mathrm{T}}^{\ell} - E_{\mathrm{T}}^{\mathrm{miss}})/p_{\mathrm{T}}^{V}$					✓	
$m_{\ell\ell}$			✓			
$\cos\theta^{*}(\vec{\ell^{-}}, \vec{V})$			✓			✓

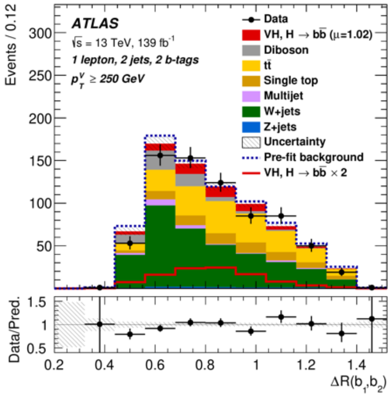
The key elements for the improvements – ML techniques

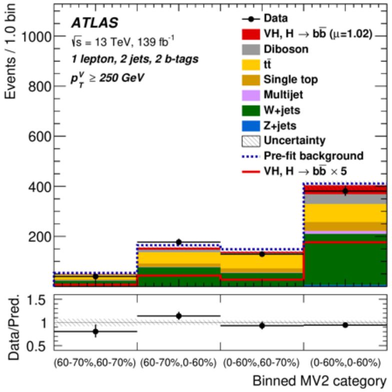
- ➢ Improving the signal and background separation

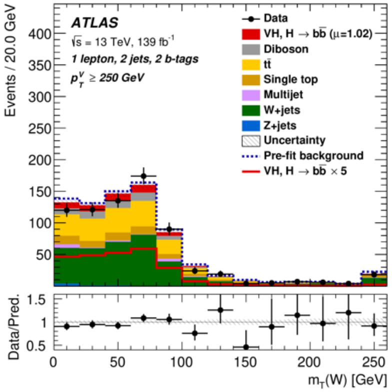
- ● Kinematic variables, some specific to 3-jet regions.
- ● m_{bb}, $\triangle R_{bb}$, p_T^V most important ones.





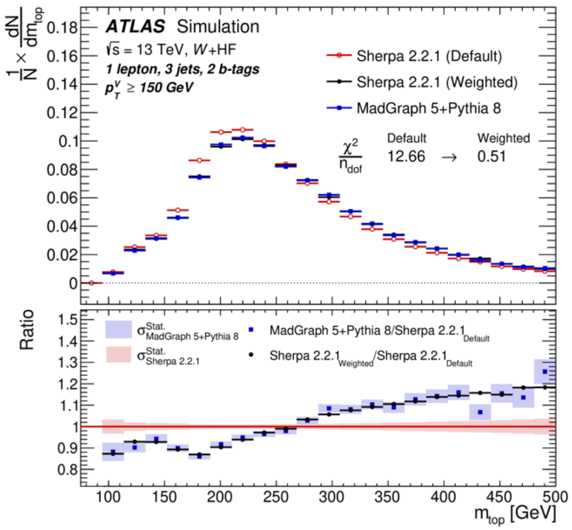


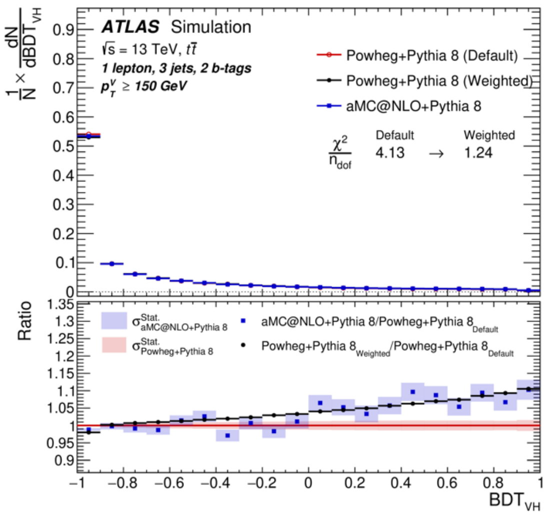




The key elements for the improvements – ML techniques

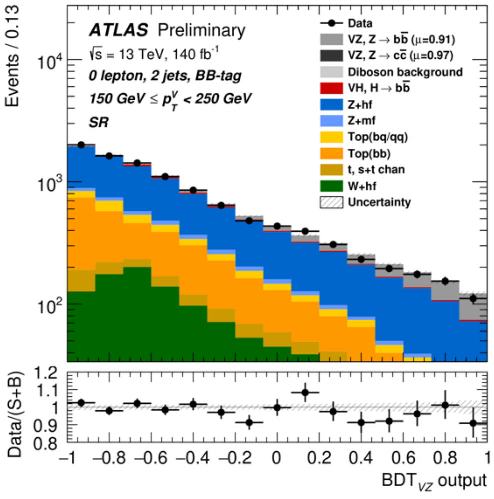
- Access the uncertainties in the modelling of the shape of backgrounds: compare the ***nominal*** and ***alternative*** MC generators
  - alternative samples: low statistics, causing issues with fit stability
  - Re-weight nominal distribution to minic the alternative samples -> take advantage of the larger statistics of the nominal sample
  - use the deep neural networks based re-weighting technique, CARL
    - Simultaneously reweight multiple features, taking into account the correlations between them
    - Particular important with fitting MVA discriminant in SR

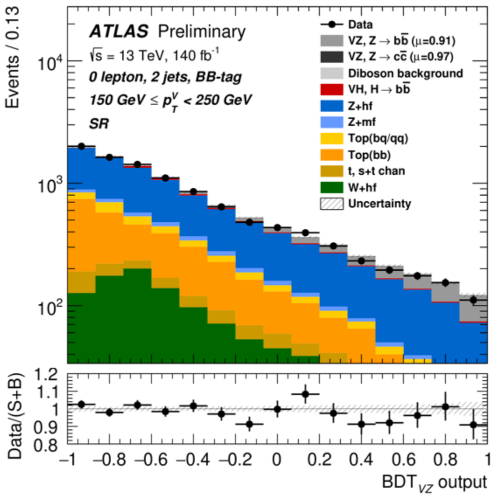


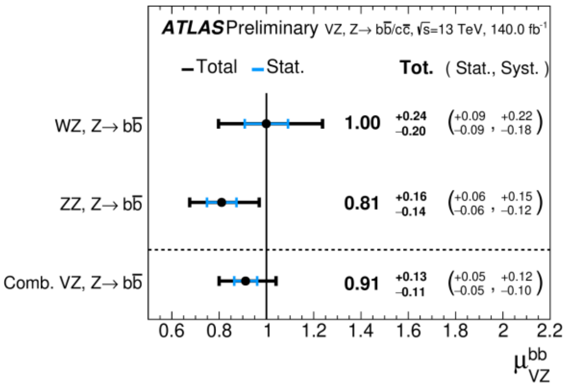


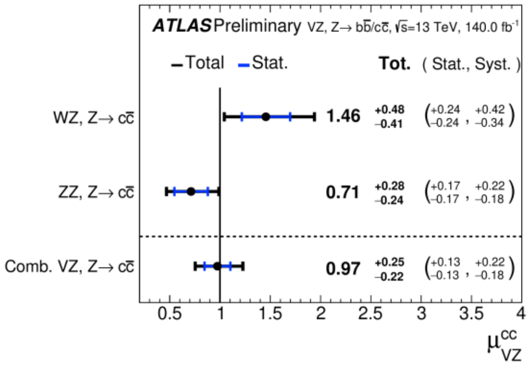
The results: the VZ diboson standard candle

- Use the VZ(bb/cc) diboson as a standard candle before checking the VH results
  - Train the BDT with VZ(bb/cc) as signal, using the exactly same set up (training parameters, inputs variables) as the VH results
  - VZ(bb) and VZ(cc) are extracted simultaneously in the fit
- Sensitivities:
- WZ(bb): 6.4 (6.5) obs. (exp.) σ -> ***First observation!* ;** ZZ(bb) > 10 σ
- WZ(cc): 3.9 (2.7); ZZ(cc): 3.1 (4.2) -> ***First observation of VZ(cc) at 5 σ with ATLAS***





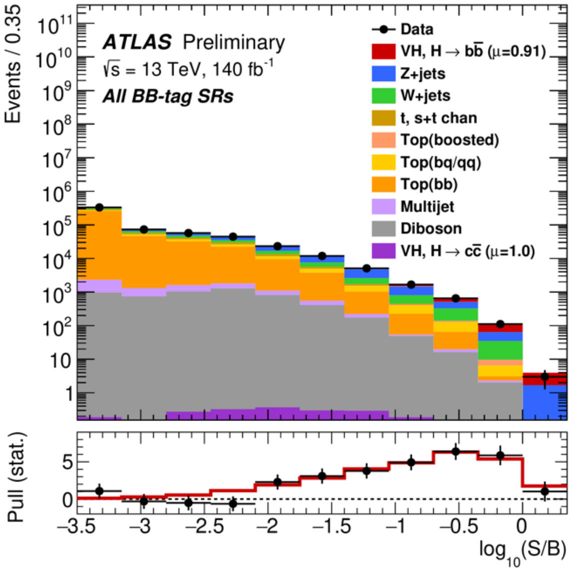




- ***Results in very good agreement with the SM predication***

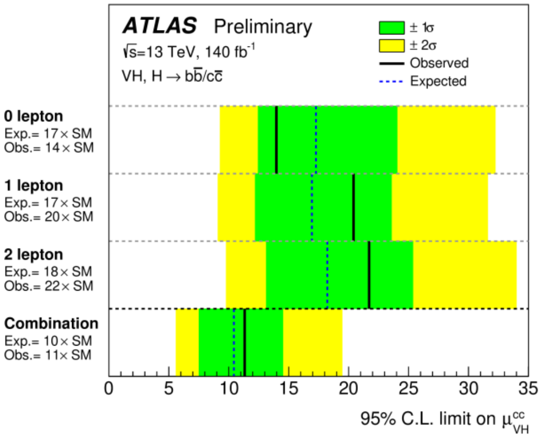
The results: VH inclusive measurement

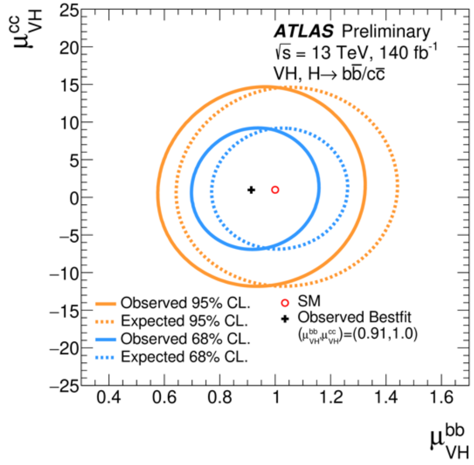
- VH(bb) and VH(cc) are extracted simultaneously in the fit
  - VHbb precison around 15%;
  - First time WH(bb) observation at 5.3 (5.5) obs. (exp.) sensitivity
  - VHcc limit at 95% CL 11.2 (10.4) obs. (exp.). Strongest observed limit to date

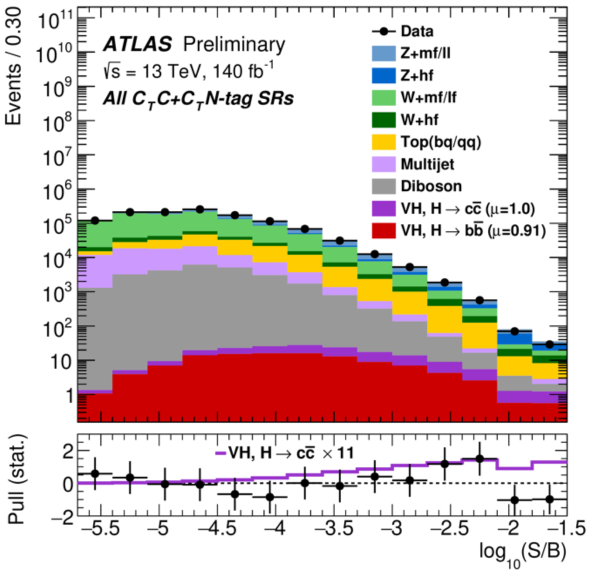


$$\mu_{VH}^{bb} = 0.91^{+0.16}_{-0.14} = 0.91 \pm 0.10\,(\text{stat.})^{+0.12}_{-0.11}\,(\text{syst.})$$

$$\mu_{VH}^{cc} = 1.0^{+5.4}_{-5.2} = 1.0^{+4.0}_{-3.9}\,(\text{stat.})^{+3.6}_{-3.5}\,(\text{syst.}).$$



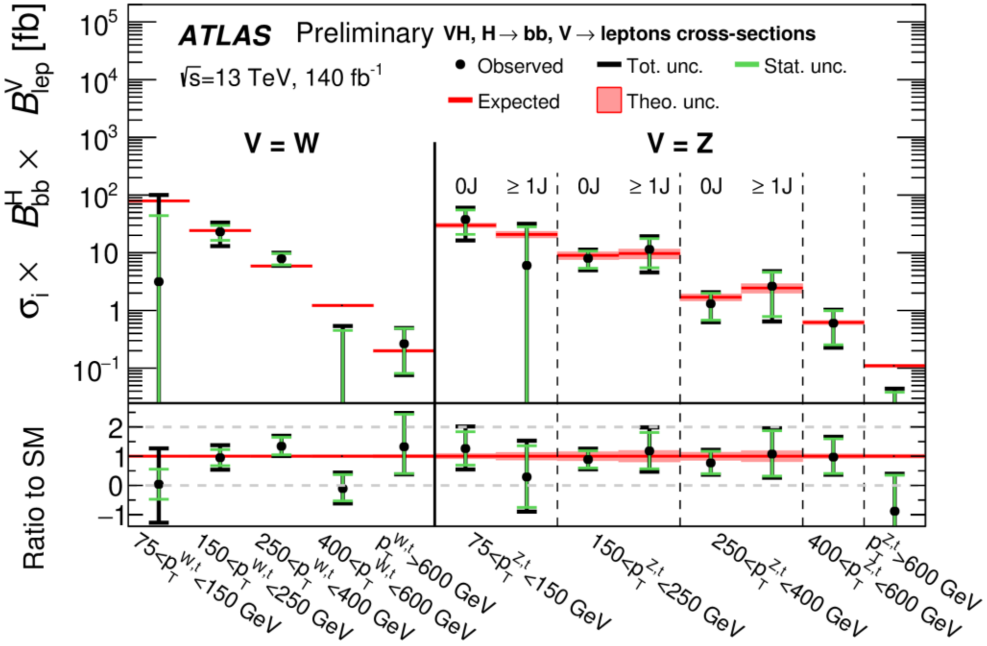




- ***Results in very good agreement with the SM predication***

The results: Differential XSec (STXS) measurement

- For the first time:
  - Extended the measurement to the new physics more sensitivity pTV > 600 GeV bin
  - Extended the ZH into different nJet bins
  - Precision in different bins vary from 35% to 100%



- ***Results in very good agreement with the SM predication, SM compatibility 90%***

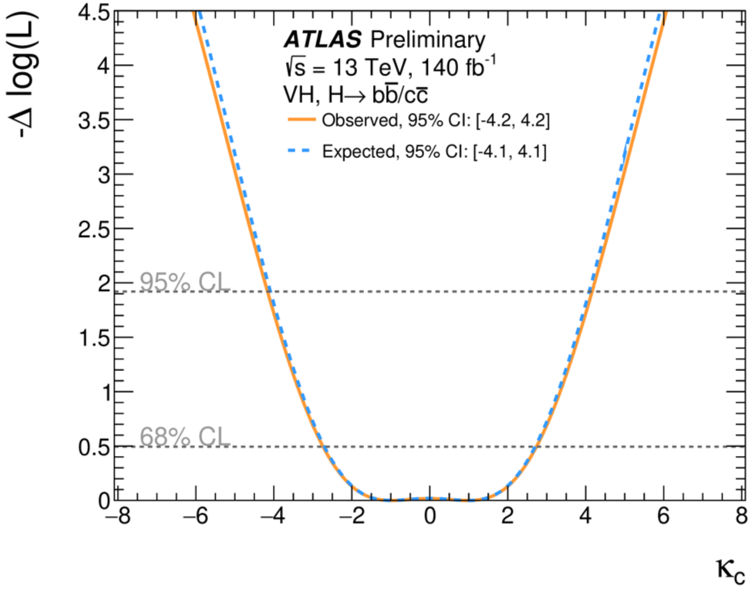
The results: Kappa interpretation

- Reparameterising the mu_VHbb and mu_VHcc in terms of the Higgs-bottom and Higgs-charm multiplicative coupling modifiers, κb and κc, others set to SM�

- 1D scan for κc, fixing κb at 1
  - | κc | < 4.2 (was < 8.5 in the previous results)
  - An equivalent approach for κb yields an observed 95% CL interval of 0.67 < | κb | < 1.38

$$\mu_{VH}^{bb} = \frac{\kappa_b^2}{1 + B_{hbb}^{SM}(\kappa_b^2 - 1) + B_{hcc}^{SM}(\kappa_c^2 - 1)}$$

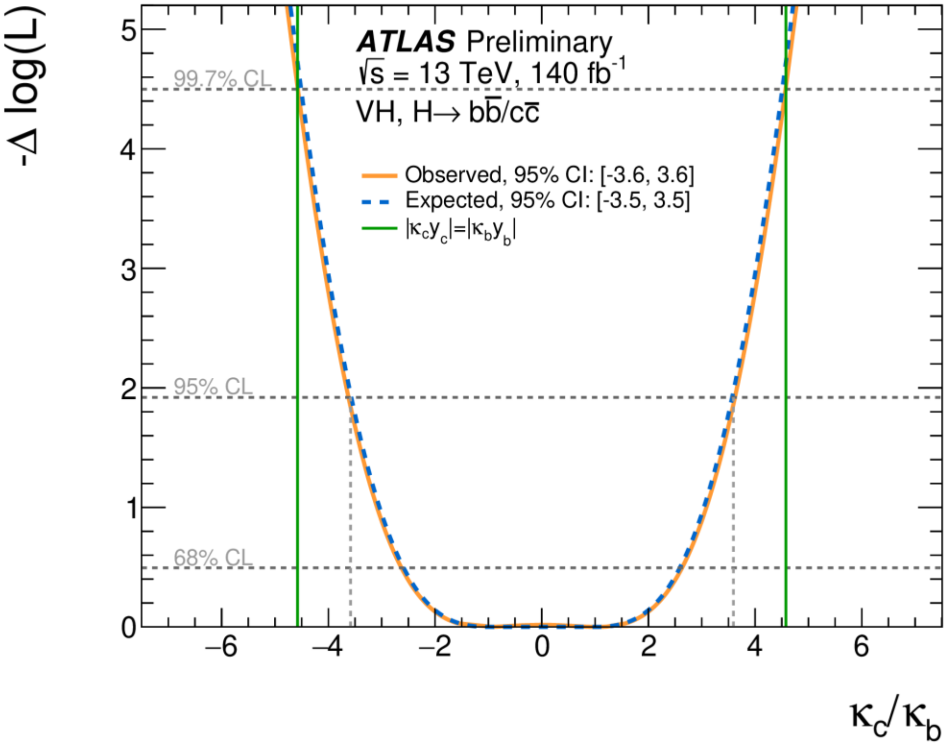
$$\mu_{VH}^{cc} = \frac{\kappa_c^2}{1 + B_{hbb}^{SM}(\kappa_b^2 - 1) + B_{hcc}^{SM}(\kappa_c^2 - 1)}$$



The results: Kappa interpretation

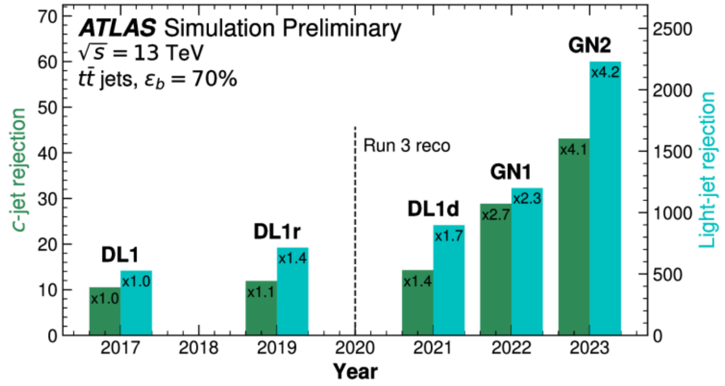
➢ Alternative parameterisation is performed targeting the ratio $\kappa c / \kappa b$

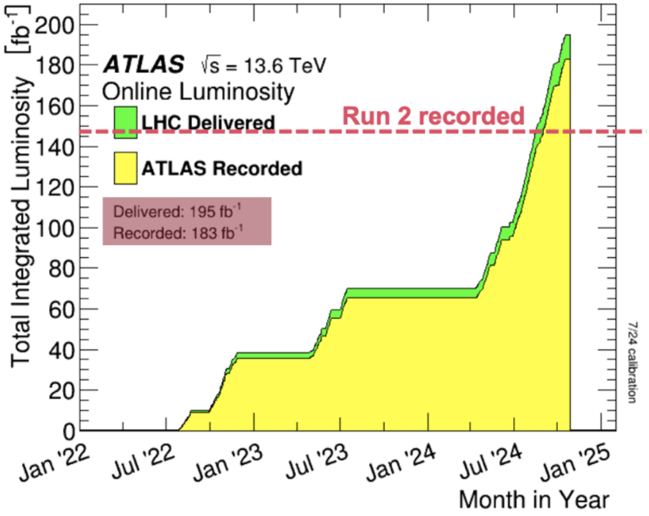
- The green vertical lines→ values of | $\kappa c / \kappa b$ | for which the Higgs-charm and Higgs-bottom couplings are equal. i.e. $\kappa c / \kappa b = mb / mc$ = 4.578+-0.008
- The coupling of the Higgs boson to charm quarks is weaker than the coupling of the Higgs boson to bottom quarks is confirmed at 3σ. �



In the near future

- Better object performance, better physics!
  - New “all in one” b/c-tagger GN available now
    - 2-4 times better performance expected
- More data more fun
  - Run 3 data taking is on-going smoothly, expected another more than 300 fb^{-1} data to come
  - Hbb: precision measurements paving the way to new physics
  - Hcc: Run 3 hot topic, further pushing the limit to new record
    - including the other production mode in the game (ggF, VBF and ttH)

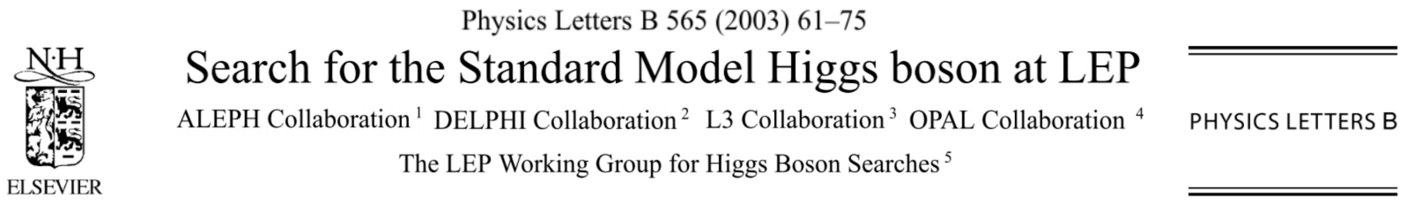




Back Up

A long way to reach the Hbb discovery

First H->bb searches started at LEP



Physics Letters B 565 (2003) 61–75

Search for the Standard Model Higgs boson at LEP

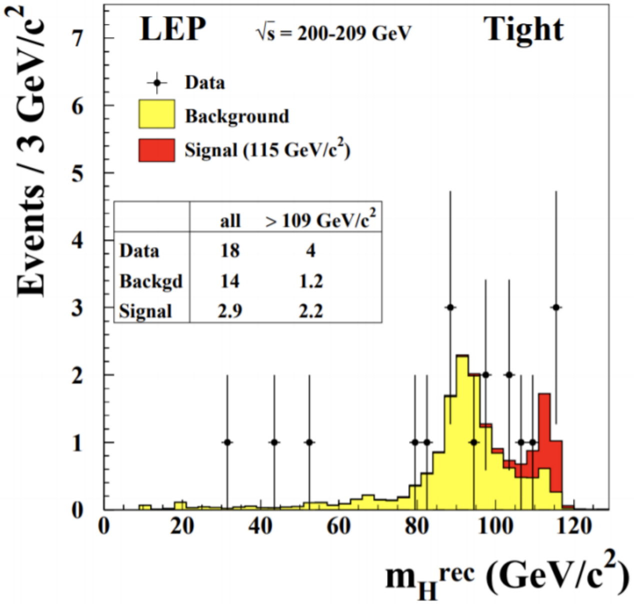
ALEPH Collaboration[1] DELPHI Collaboration[2] L3 Collaboration[3] OPAL Collaboration[4]

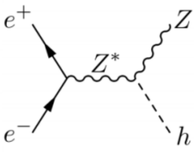
The LEP Working Group for Higgs Boson Searches[5]

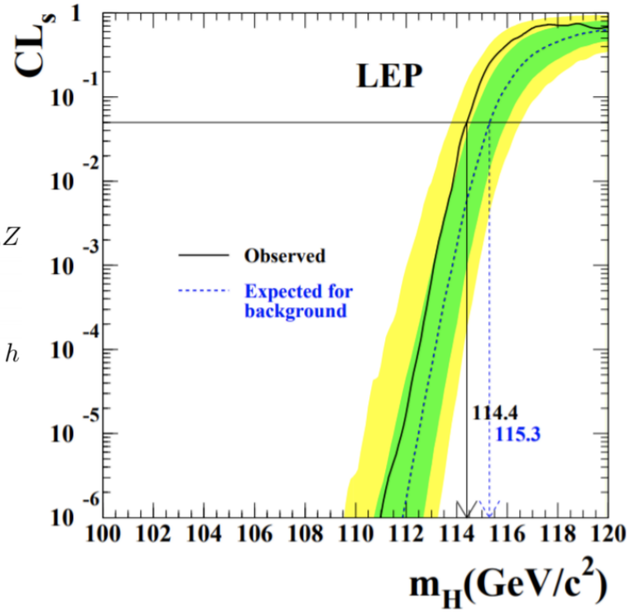
ELSEVIER

PHYSICS LETTERS B

$m_H > 114.4$ GeV @ 95%CL







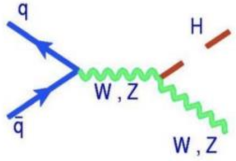
A long way to reach the Hbb discovery

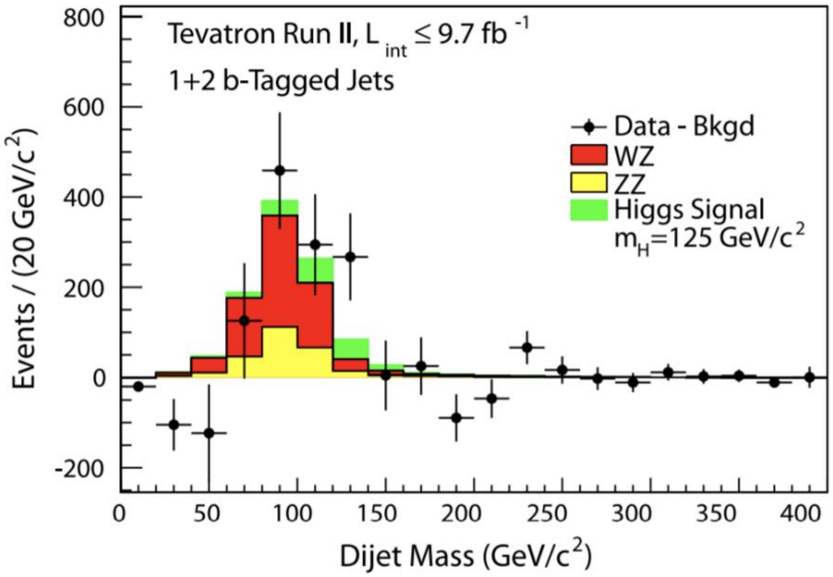
…and continued at Tevatron by focusing on the VH production…

PRL **109,** 071804 (2012) PHYSICAL REVIEW LETTERS week ending 17 AUGUST 2012

Evidence for a Particle Produced in Association with Weak Bosons and Decaying to a Bottom-Antibottom Quark Pair in Higgs Boson Searches at the Tevatron

(*CDF Collaboration)
(†D0 Collaboration)





<u>Significance</u>

2.8σ observed @ 125 GeV

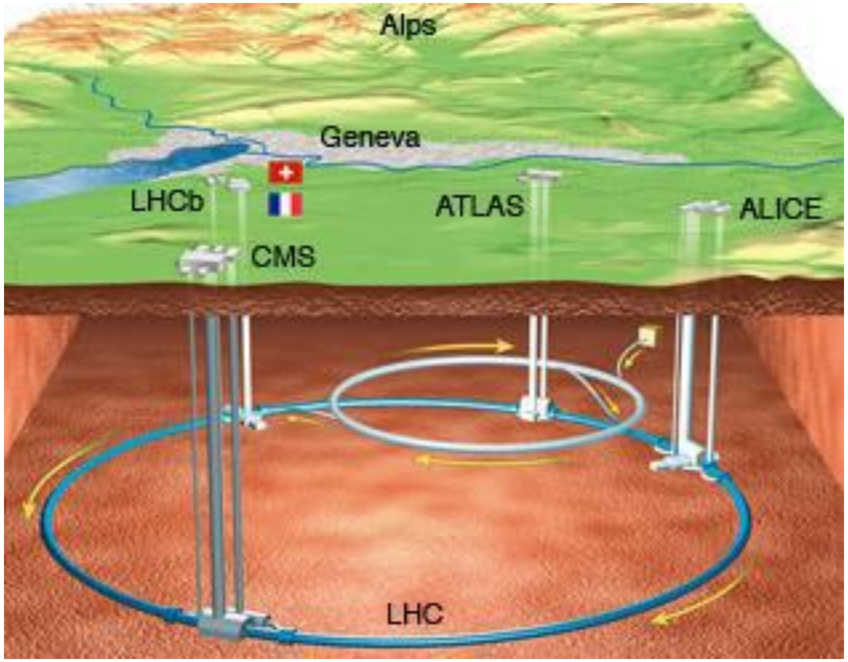
A long way to reach the Hbb discovery

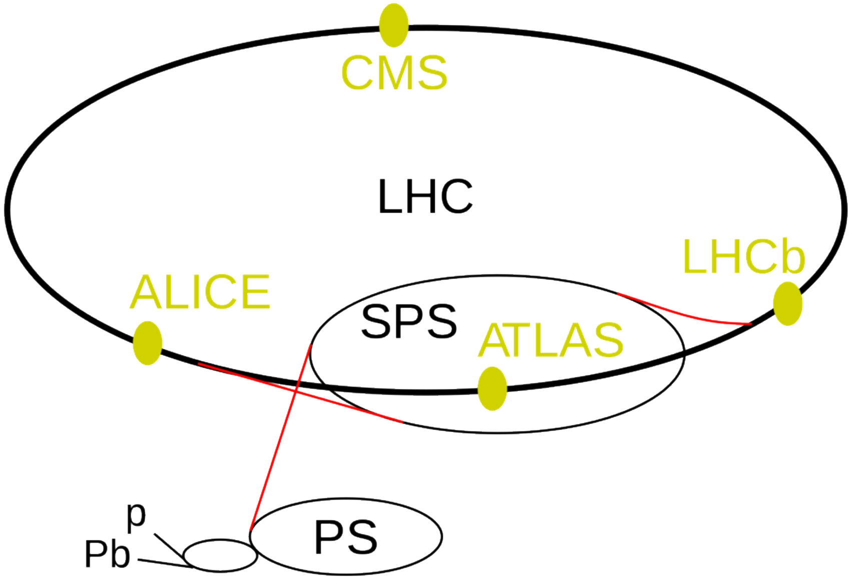
…and continued at LHC by focusing on the VH production…

	Signal strength	Significance (exp)	Significance (obs)	
ATLAS Run 1	0.52+-0.38	2.6	1.4	
CMS Run 1	0.89+-0.45	2.5	2.1	
LHC Run 1 combination	0.79+-0.28	3.7	2.6	
ATLAS Run2 2015-2016 (36fb-1)	1.20+-0.39	3.0	***3.5***	***First evidence***
CMS Run2 2015-2016 (36fb-1)	1.19+-0.39	2.8	***3.3***	***First evidence***
ATLAS Run2 2015-2017 (80fb-1)	1.16+-0.26	4.3	***4.9***	**Observation at 5.4 (5.5) when combining with all the other Hbb searches (VHbb Run 1,ttH,VBF)**
CMS Run2 2015-2017 (77fb-1)	1.01+-0.23	4.8	***4.9***	**Observation at 5.6 (5.5) when combining with all the other Hbb searches (VHbb Run 1,ttH,VBF)**

The LHC collider

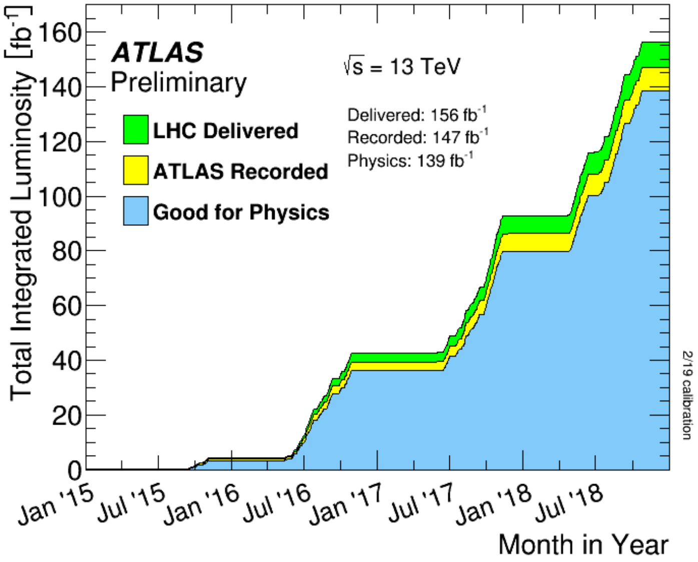
- The largest and highest-energy particle collider in the world.
- Housed in a circular tunnel with 27 km in circumference and 45-175 m in depth underground.

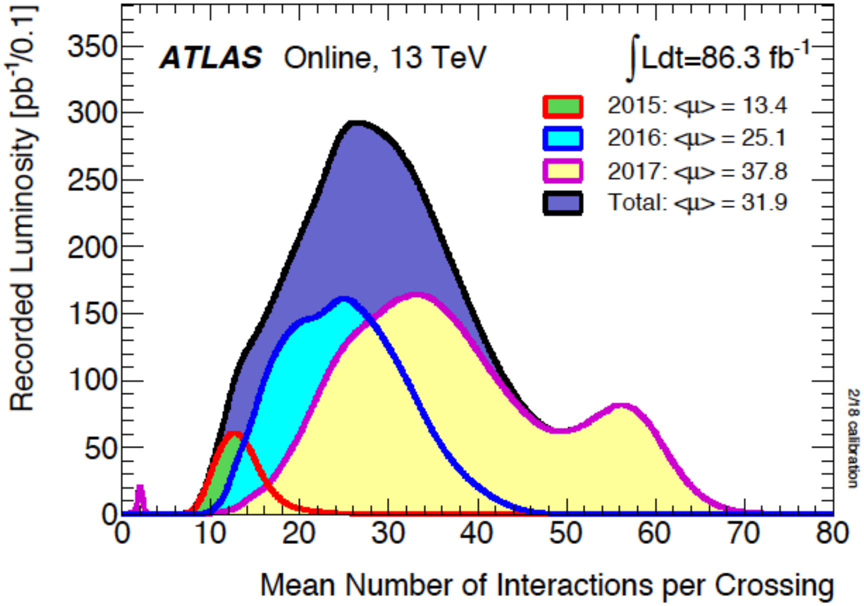




- Four main experiments: ATLAS, CMS, LHCb, ALICE.
- Designed proton-proton collision energy: 14 TeV (13 TeV at Run 2).

LHC performance





- ➢ Stunning performance of the LHC: lumi up to 2 $*10^{34}$ cm^{-2} s^{-1}
- ➢ Excellent operation of the ATLAS detector
- ➢ High rates and large pile-up: Challenges for triggers, jets reconstruction, b-tagging…

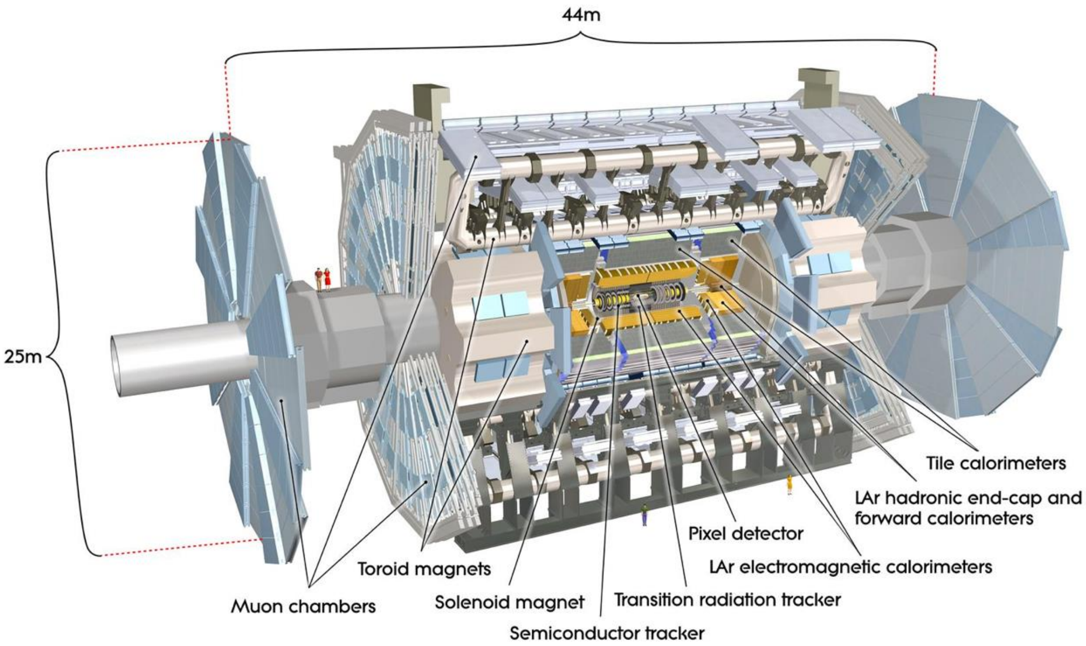
The ATLAS detector

- World's largest particle detector with a diameter of 25 m and length of 44 m.

- General-purpose detector, designed mainly to search for the Higgs boson and new physics.

- Sub-detectors:

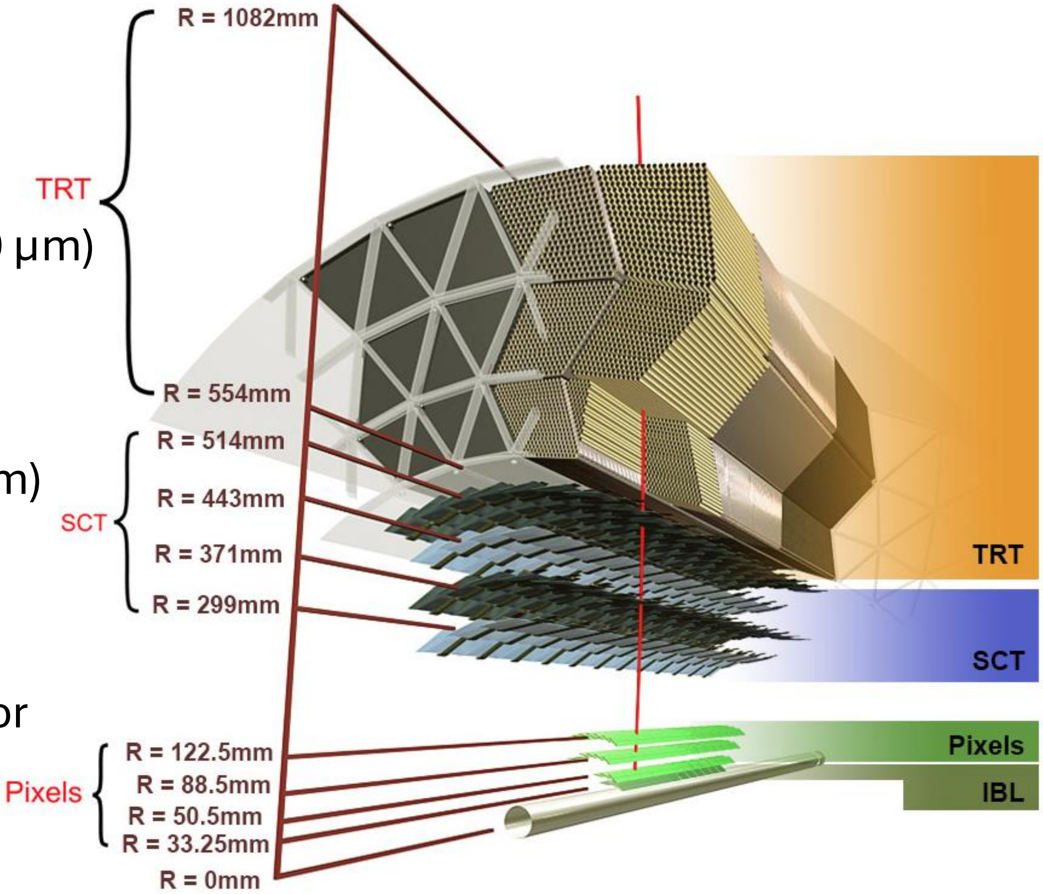
  - **Inner detector**:
    - Measure the trajectories and momenta of charged particles
  - **EM and Hadronic calorimeter**
    - Measure the energy of electrons, photons and hadrons
  - **Muon spectrometer**
    - Measure the trajectories and momenta of muon

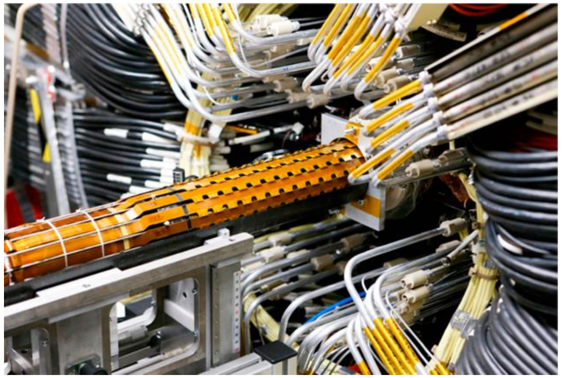


The ATLAS detector---ID and IBL

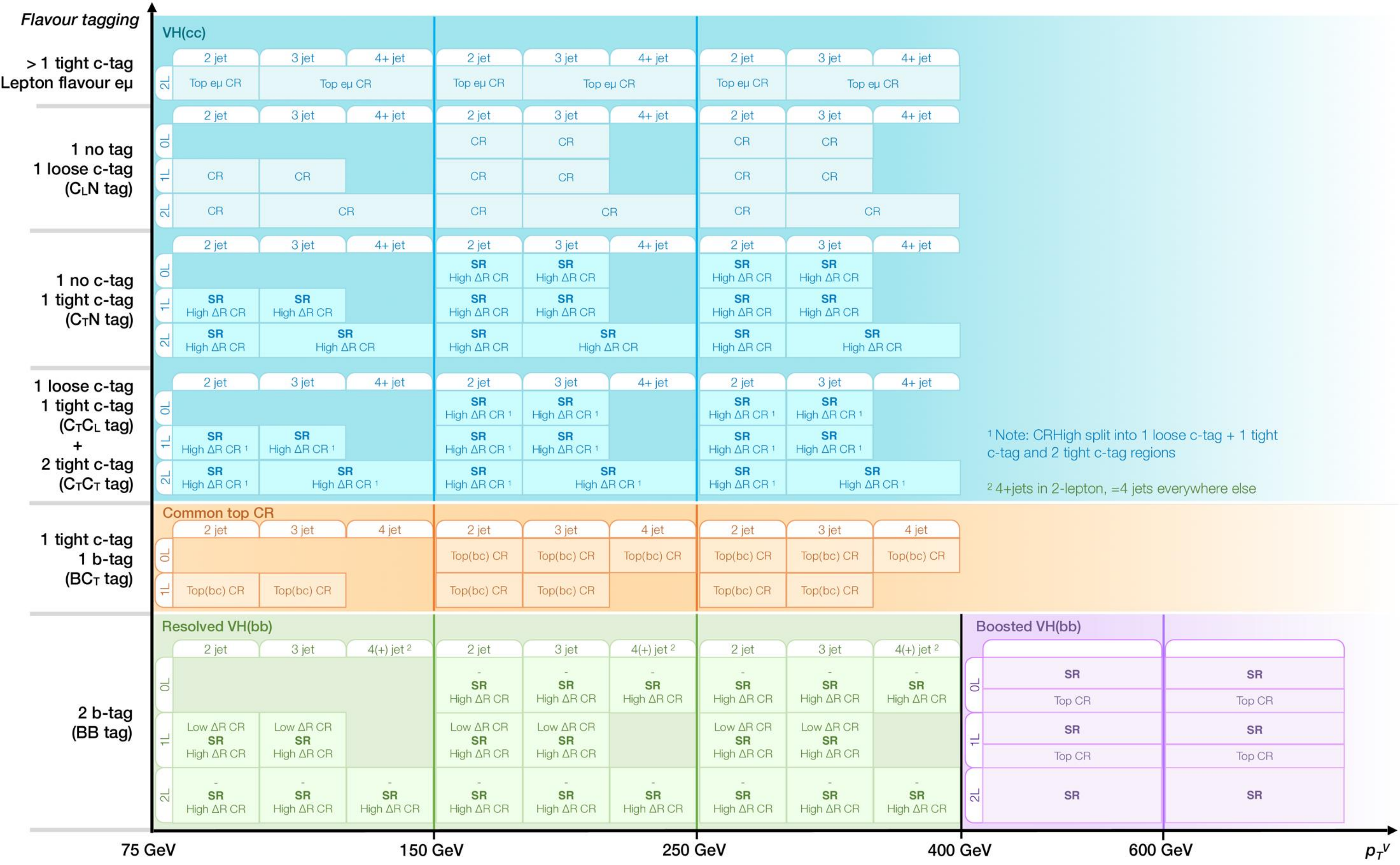
- IBL

- Smaller pixel size (50x250 vs 50x400 μm)
- Closer to interaction region (R~3.3cm)
- H → bb primary physics motivation for the new detector!
- Improvement of 10% for the b-tagging algorithm performance in Run 2





More detailed look for all the fit regions



VH(cc)

Flavour tagging		75 GeV			150 GeV			250 GeV		
		2 jet	3 jet	4+ jet	2 jet	3 jet	4+ jet	2 jet	3 jet	4+ jet
> 1 tight c-tag Lepton flavour eμ	2L	Top eμ CR	Top eμ CR		Top eμ CR	Top eμ CR		Top eμ CR	Top eμ CR	
1 no tag 1 loose c-tag (C_LN tag)	0L				CR	CR		CR	CR	
	1L	CR	CR		CR	CR		CR	CR	
	2L	CR	CR		CR	CR		CR	CR	
1 no c-tag 1 tight c-tag (C_TN tag)	0L				**SR** High ΔR CR	**SR** High ΔR CR		**SR** High ΔR CR	**SR** High ΔR CR	
	1L	**SR** High ΔR CR	**SR** High ΔR CR		**SR** High ΔR CR	**SR** High ΔR CR		**SR** High ΔR CR	**SR** High ΔR CR	
	2L	**SR** High ΔR CR	**SR** High ΔR CR		**SR** High ΔR CR	**SR** High ΔR CR		**SR** High ΔR CR	**SR** High ΔR CR	
1 loose c-tag 1 tight c-tag (C_TC_L tag) + 2 tight c-tag (C_TC_T tag)	0L				**SR** High ΔR CR[1]	**SR** High ΔR CR[1]		**SR** High ΔR CR[1]	**SR** High ΔR CR[1]	
	1L	**SR** High ΔR CR[1]	**SR** High ΔR CR[1]		**SR** High ΔR CR[1]	**SR** High ΔR CR[1]		**SR** High ΔR CR[1]	**SR** High ΔR CR[1]	
	2L	**SR** High ΔR CR[1]	**SR** High ΔR CR[1]		**SR** High ΔR CR[1]	**SR** High ΔR CR[1]		**SR** High ΔR CR[1]	**SR** High ΔR CR[1]	

Common top CR

Flavour tagging		75 GeV			150 GeV			250 GeV		
		2 jet	3 jet	4 jet	2 jet	3 jet	4 jet	2 jet	3 jet	4 jet
1 tight c-tag 1 b-tag (BC_T tag)	0L				Top(bc) CR	Top(bc) CR	Top(bc) CR	Top(bc) CR	Top(bc) CR	Top(bc) CR
	1L	Top(bc) CR	Top(bc) CR		Top(bc) CR	Top(bc) CR		Top(bc) CR	Top(bc) CR	

Resolved VH(bb)

Flavour tagging		75 GeV			150 GeV			250 GeV		
		2 jet	3 jet	4(+) jet [2]	2 jet	3 jet	4(+) jet [2]	2 jet	3 jet	4(+) jet [2]
2 b-tag (BB tag)	0L				- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR
	1L	Low ΔR CR **SR** High ΔR CR	Low ΔR CR **SR** High ΔR CR		Low ΔR CR **SR** High ΔR CR	Low ΔR CR **SR** High ΔR CR		Low ΔR CR **SR** High ΔR CR	Low ΔR CR **SR** High ΔR CR	
	2L	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR	- **SR** High ΔR CR

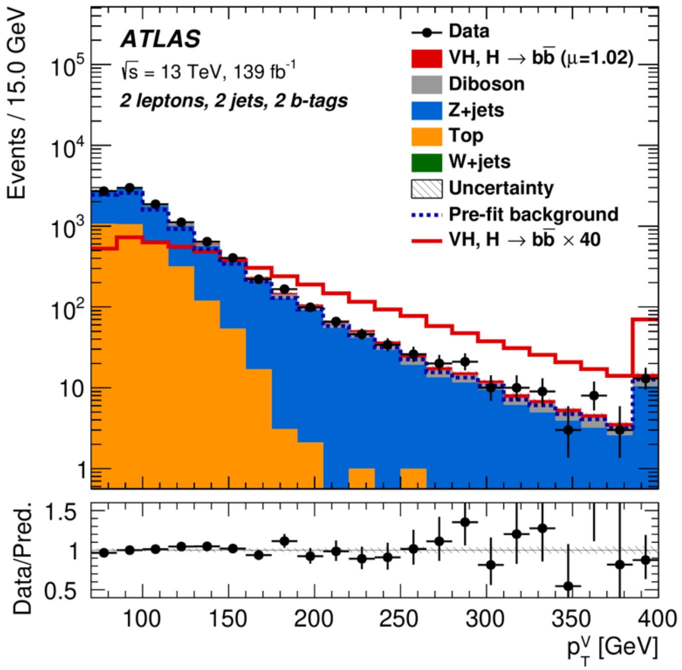
Boosted VH(bb)

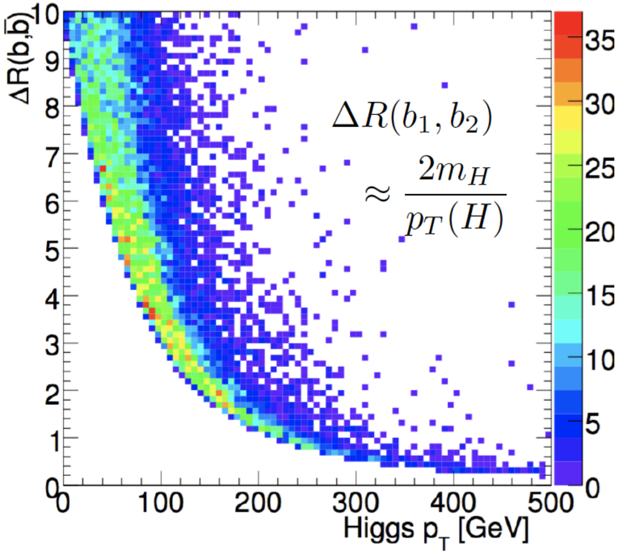
Flavour tagging		400 GeV	600 GeV
2 b-tag (BB tag)	0L	**SR** Top CR	**SR** Top CR
	1L	**SR** Top CR	**SR** Top CR
	2L	**SR**	**SR**

[1] Note: CRHigh split into 1 loose c-tag + 1 tight c-tag and 2 tight c-tag regions

[2] 4+jets in 2-lepton, =4 jets everywhere else

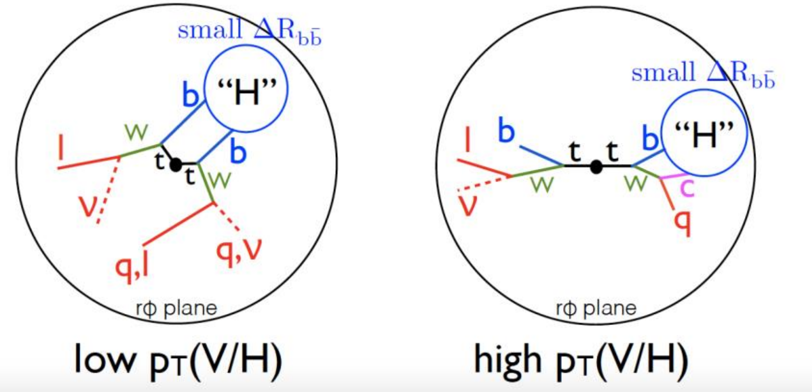
The "high pT" regime





- Requiring high pT(V) (or pT(H)) suppresses background significantly more than signal, improving S/B ratio
  - Used for event classification
  - Added as input variable in MVA analysis
- H→bb a simple 2-body decay, can requie low dR(bb) at high pT(H), with almost no loss in signal efficiency

Ttbar background



- Backgrounds (esp. ttbar) significantly suppressed by these requirements

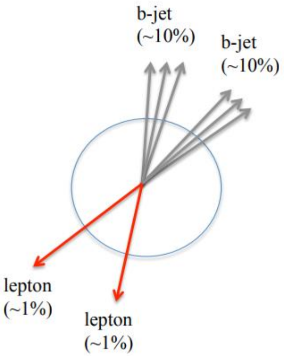
The key ingredients for the improvements – mass resolution

- Sharpening signal mass peak directly improves sensitivity

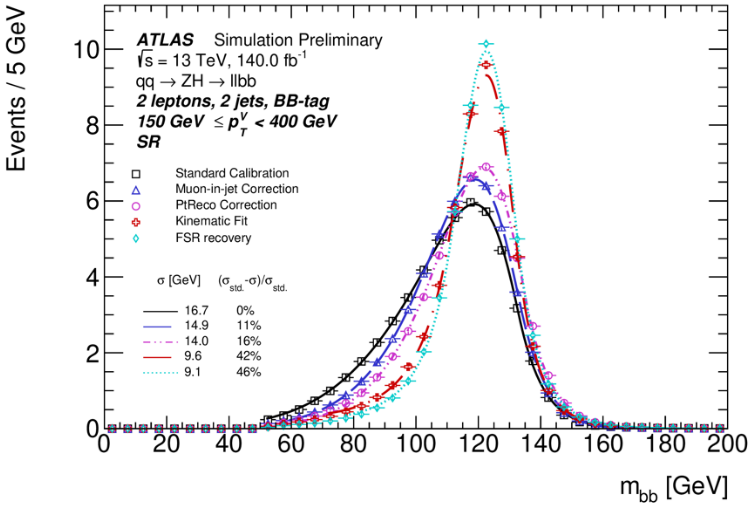
- Two most important correction

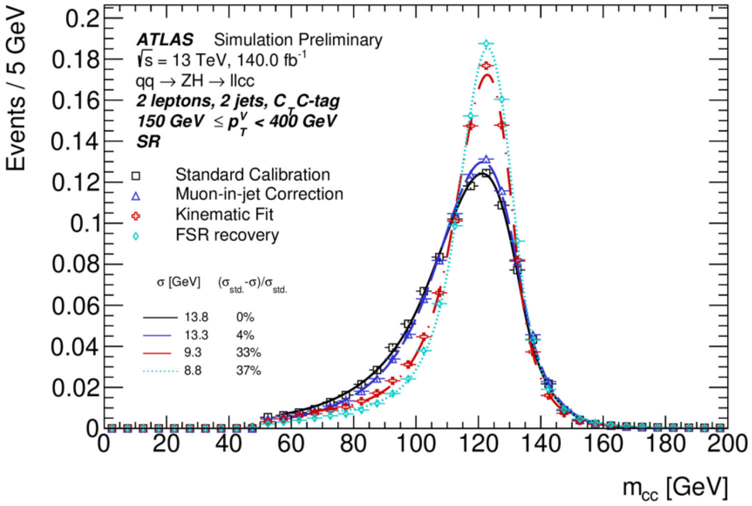
- μ-in-jet correction: if available, add muon to jet momentum (+13%)

- For 2-lepton channel (≤3 jets), use full kinematic likelihood fit



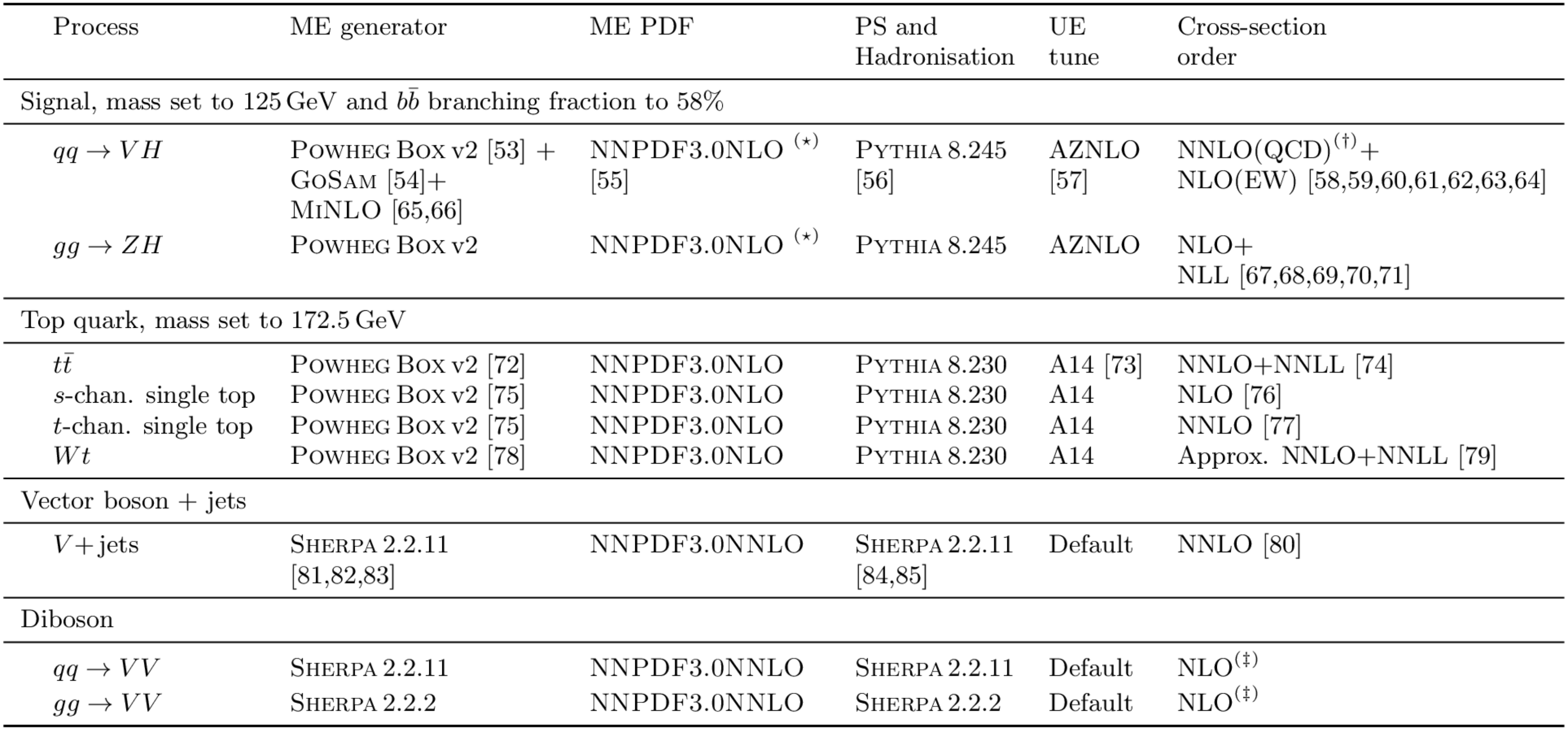
  - Final state fully reconstructed
  - High resolution on leptons
  - Mass resolution improvement: 40%





- Final state radiation (FSR) recovery further minimising the effect of hard QCD radiation of heavy quarks in the Higgs boson reconstruction.

Signal and Backgrounds Samples



Process	ME generator	ME PDF	PS and Hadronisation	UE tune	Cross-section order
Signal, mass set to 125 GeV and $b\bar{b}$ branching fraction to 58%					
$qq \to VH$	POWHEG BOX v2 [53] + GOSAM [54]+ MINLO [65,66]	NNPDF3.0NLO $^{(\star)}$ [55]	PYTHIA 8.245 [56]	AZNLO [57]	NNLO(QCD)$^{(\dagger)}$+ NLO(EW) [58,59,60,61,62,63,64]
$gg \to ZH$	POWHEG BOX v2	NNPDF3.0NLO $^{(\star)}$	PYTHIA 8.245	AZNLO	NLO+ NLL [67,68,69,70,71]
Top quark, mass set to 172.5 GeV					
$t\bar{t}$	POWHEG BOX v2 [72]	NNPDF3.0NLO	PYTHIA 8.230	A14 [73]	NNLO+NNLL [74]
s-chan. single top	POWHEG BOX v2 [75]	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO [76]
t-chan. single top	POWHEG BOX v2 [75]	NNPDF3.0NLO	PYTHIA 8.230	A14	NNLO [77]
Wt	POWHEG BOX v2 [78]	NNPDF3.0NLO	PYTHIA 8.230	A14	Approx. NNLO+NNLL [79]
Vector boson + jets					
V+jets	SHERPA 2.2.11 [81,82,83]	NNPDF3.0NNLO	SHERPA 2.2.11 [84,85]	Default	NNLO [80]
Diboson					
$qq \to VV$	SHERPA 2.2.11	NNPDF3.0NNLO	SHERPA 2.2.11	Default	NLO$^{(\ddagger)}$
$gg \to VV$	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	Default	NLO$^{(\ddagger)}$

Signal and Backgrounds Samples

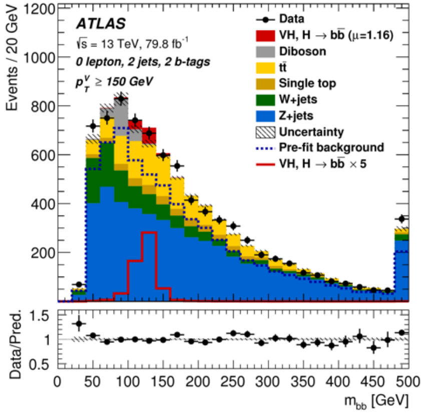
➢ Signal

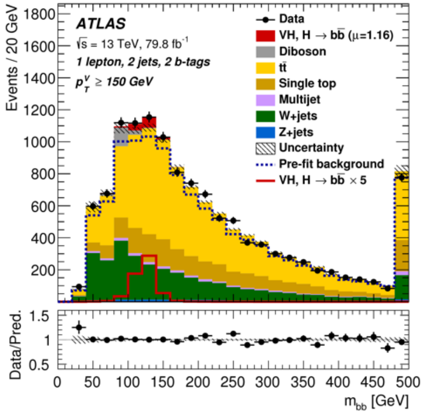
- Both qqVH and ggZH using latest Powheg+MiNLO + Pythia8 samples

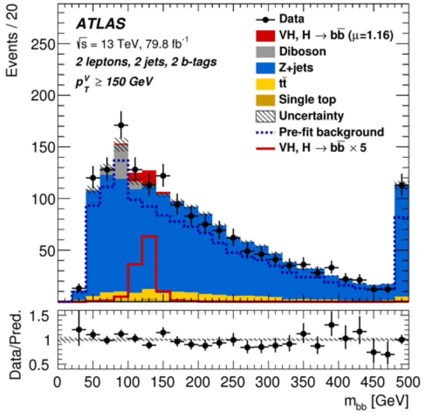
➢ Background

- V (W/Z)+jets : Sherpa 2.2.1 with jet flavor filter
- Dibson : Sherpa 2.2.1 for quark induced samples (qqVV). After EPS, include also gluon induced (ggVV) samples with Sherpa 2.2.2
- ttbar : Powheg+Pythia8, 2-lepton also incorporates di-lepton filtered sample. Dedicated MET filter ttbar samples also used in 0 lepton
- Single-top : updated to Powheg+Pythia8 samples since E
- Multijet

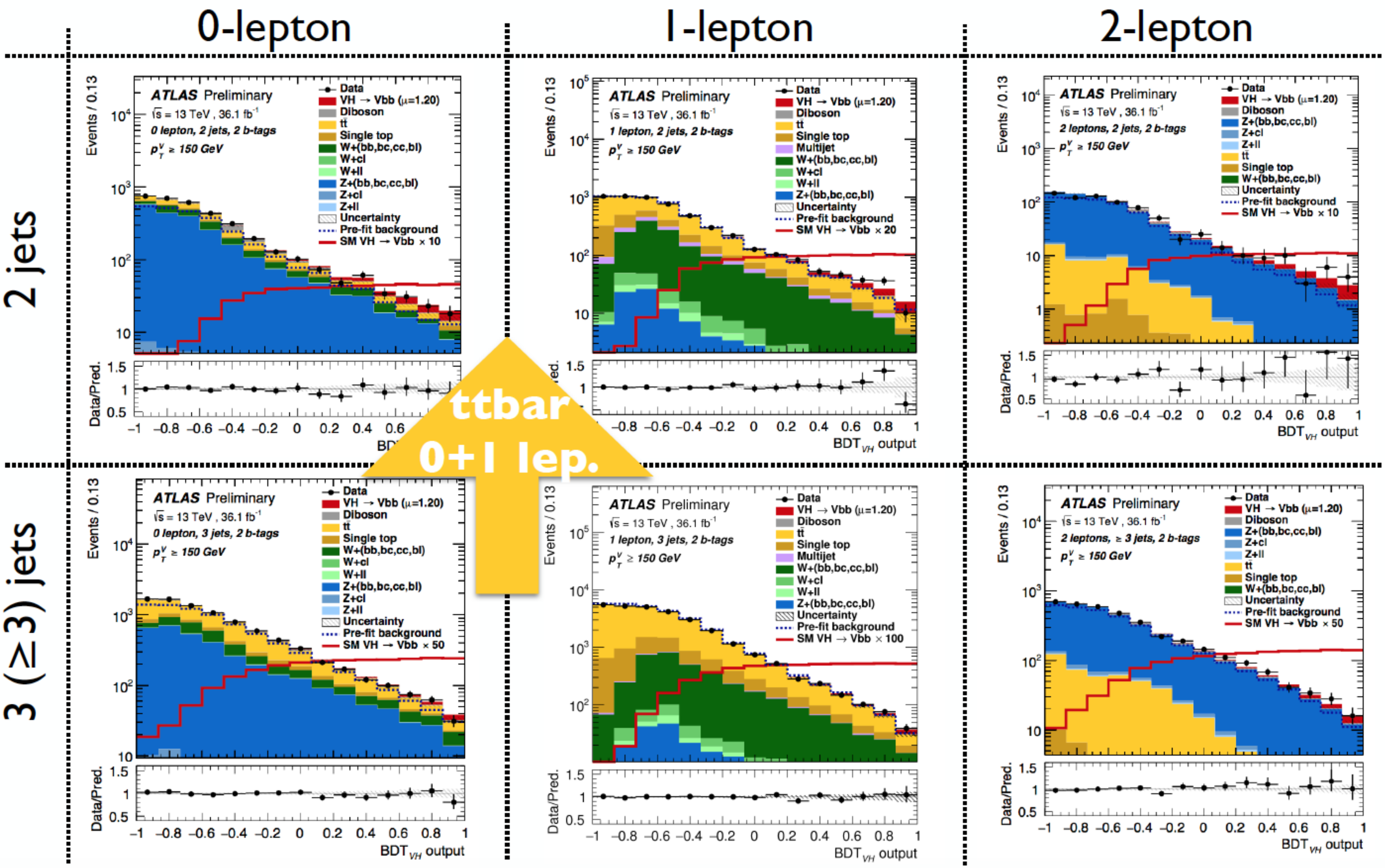
Negligible in 0 and 2 lepton (confirmed by lots of detailed studies), data-driven in 1 lepton channel (fraction: ~2-3%)





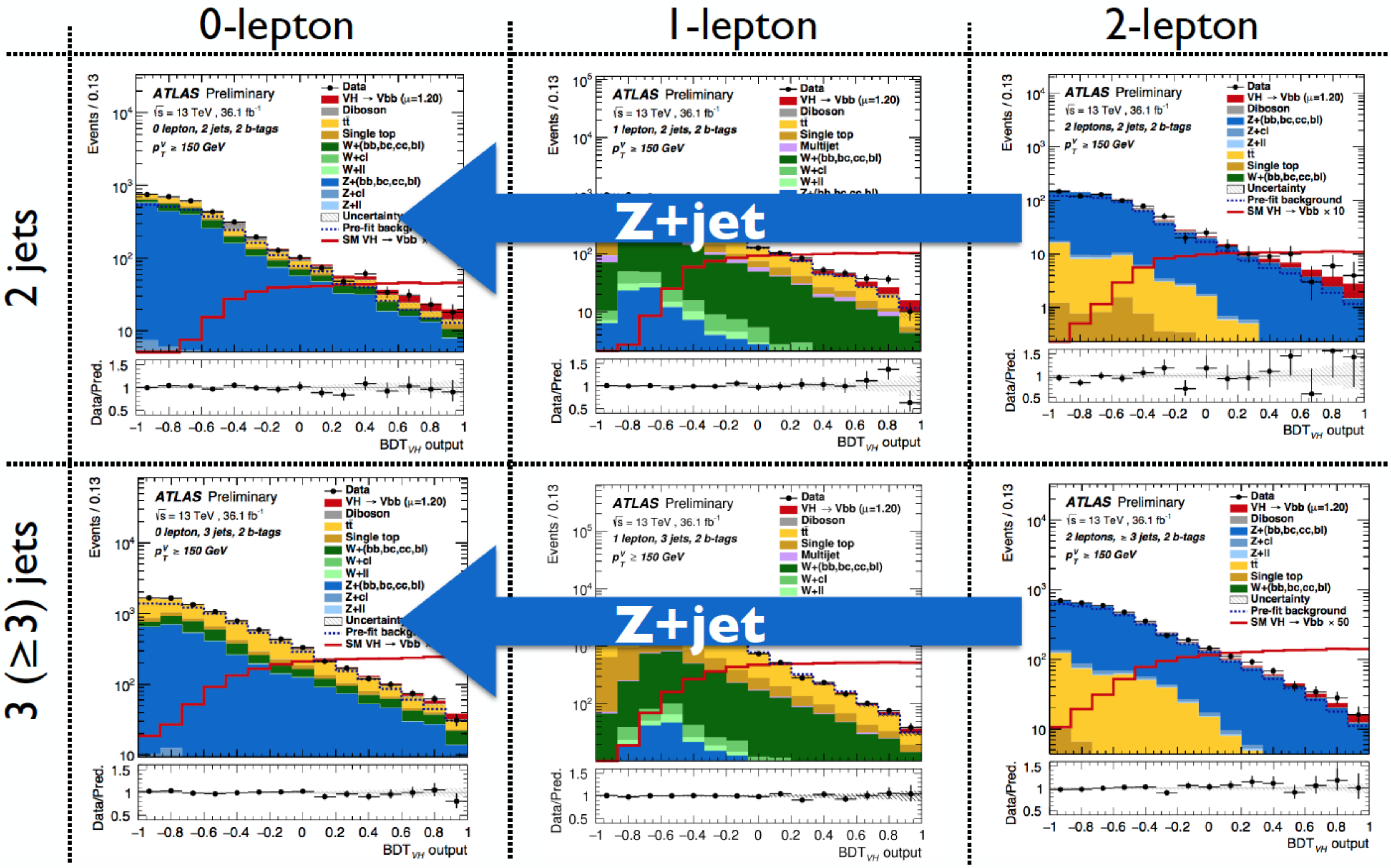


Background Modelling

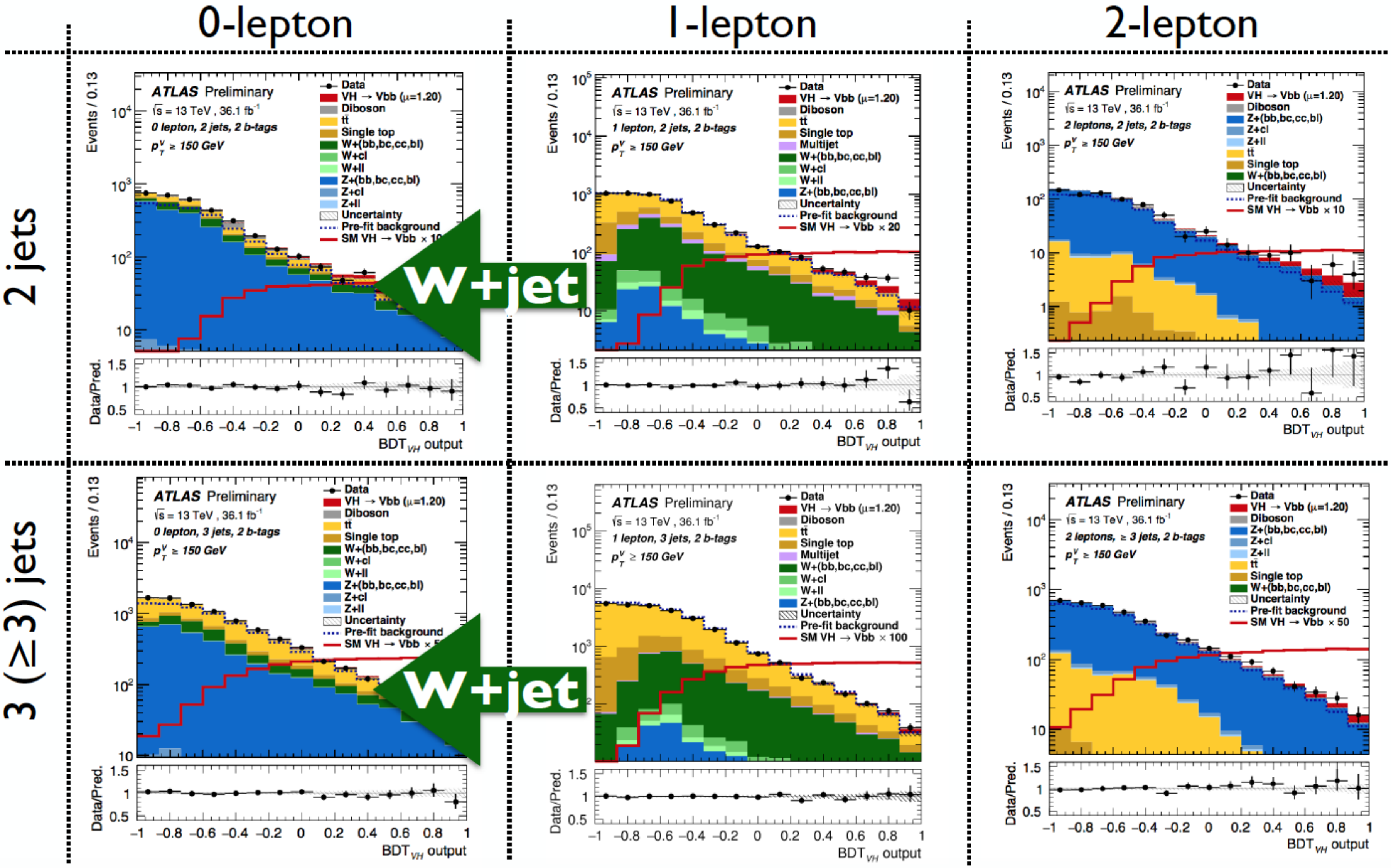


ttbar
0+1 lep.

Background Modelling



Background Modelling



0-lepton

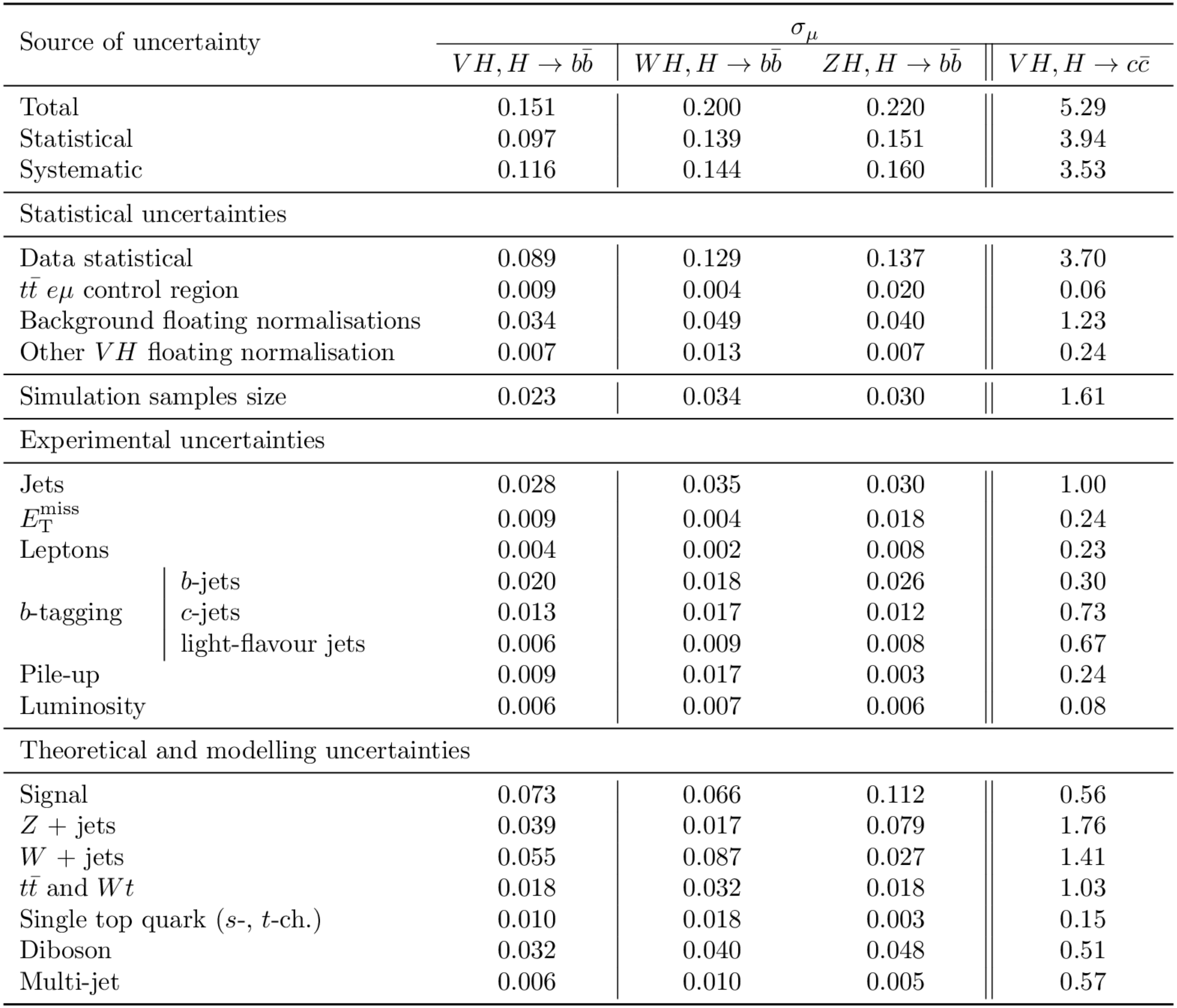
1-lepton

2-lepton

2 jets

3 (≥3) jets

Systematic breakdown

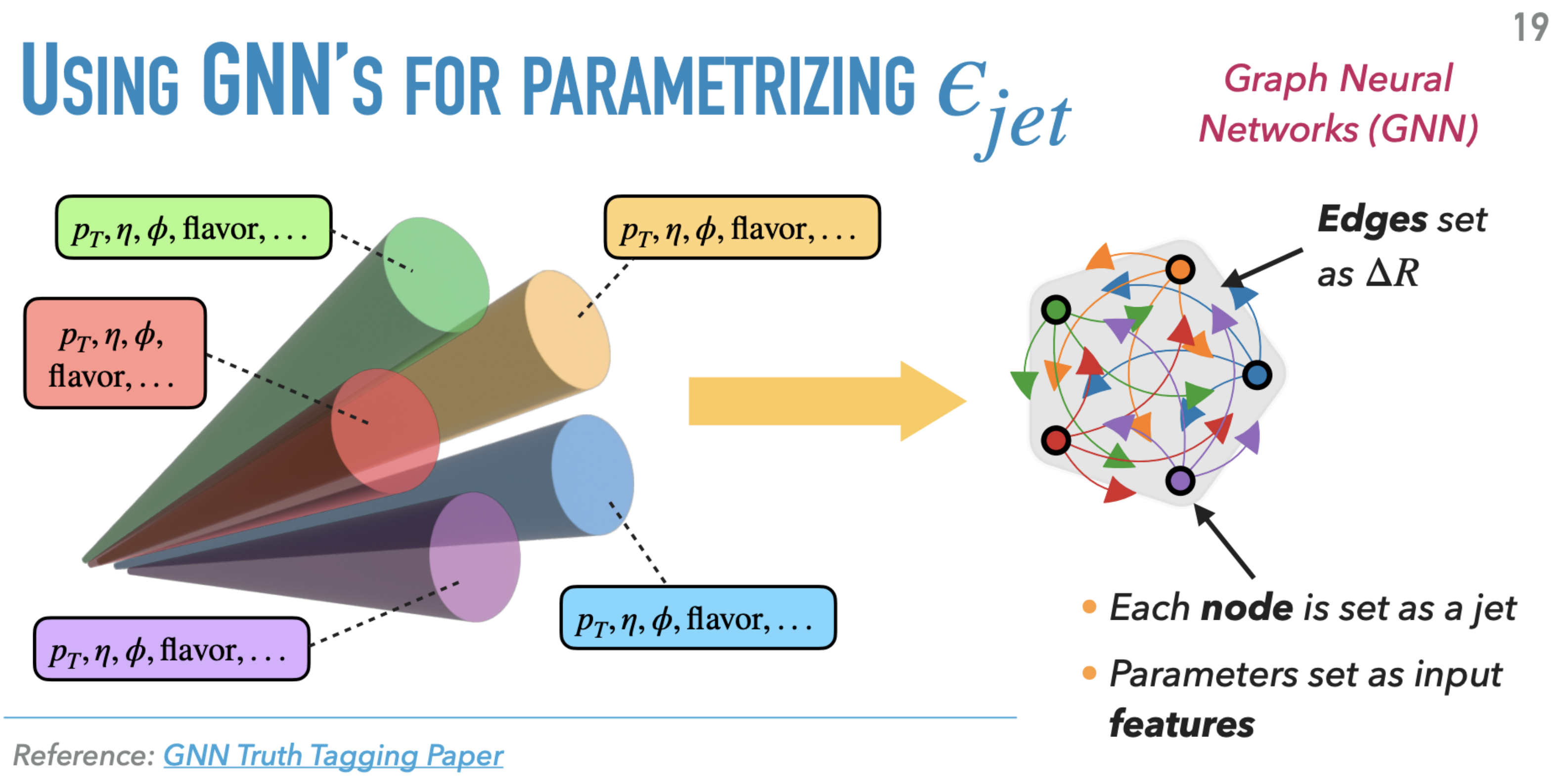


Source of uncertainty		σ_μ			
		$VH, H \to b\bar{b}$	$WH, H \to b\bar{b}$	$ZH, H \to b\bar{b}$	$VH, H \to c\bar{c}$
Total		0.151	0.200	0.220	5.29
Statistical		0.097	0.139	0.151	3.94
Systematic		0.116	0.144	0.160	3.53
Statistical uncertainties					
Data statistical		0.089	0.129	0.137	3.70
$t\bar{t}$ $e\mu$ control region		0.009	0.004	0.020	0.06
Background floating normalisations		0.034	0.049	0.040	1.23
Other VH floating normalisation		0.007	0.013	0.007	0.24
Simulation samples size		0.023	0.034	0.030	1.61
Experimental uncertainties					
Jets		0.028	0.035	0.030	1.00
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.009	0.004	0.018	0.24
Leptons		0.004	0.002	0.008	0.23
b-tagging	b-jets	0.020	0.018	0.026	0.30
	c-jets	0.013	0.017	0.012	0.73
	light-flavour jets	0.006	0.009	0.008	0.67
Pile-up		0.009	0.017	0.003	0.24
Luminosity		0.006	0.007	0.006	0.08
Theoretical and modelling uncertainties					
Signal		0.073	0.066	0.112	0.56
Z + jets		0.039	0.017	0.079	1.76
W + jets		0.055	0.087	0.027	1.41
$t\bar{t}$ and Wt		0.018	0.032	0.018	1.03
Single top quark (s-, t-ch.)		0.010	0.018	0.003	0.15
Diboson		0.032	0.040	0.048	0.51
Multi-jet		0.006	0.010	0.005	0.57

Systematic breakdown

STXS region			SM prediction	Measurement	Stat. unc.	Syst. unc. [fb]		
Process	$p_{\mathrm{T}}^{V,\mathrm{t}}$ interval	$N_{\mathrm{jet}}^{\mathrm{t}}$	[fb]	[fb]	[fb]	Th. sig.	Th. bkg.	Exp.
$W(\ell\nu)H$	75–150 GeV	≥0	79 .2 ± 2.8	3 ± 100	41	13	88	35
	150–250 GeV	≥0	24.3 ± 1.0	23 ± 10	7	2	7	3
	250–400 GeV	≥0	5.90 ± 0.25	7.9 ± 2.0	1.8	0.5	0.8	0.3
	400–600 GeV	≥0	1.03 ± 0.05	-0.11 ± 0.54	0.46	0.05	0.24	0.09
	> 600 GeV	≥0	0.20 ± 0.01	0.26 ± 0.21	0.20	0.02	0.04	0.03
$Z(\ell\ell/\nu\nu)H$	75–150 GeV	≥0	50.7 ± 3.9	51 ± 32	24	8	19	11
		=0	29.9 ± 2.5	38 ± 22	17	4	12	6
		≥1	20.7 ± 2.6	6 ± 25	25	6	9	8
	150–250 GeV	≥0	18.7 ± 2.3	18 ± 6.0	4.5	2.5	3.0	1.0
		=0	9.0. ± 1.3	8.0 ± 3.2	2.7	0.9	1.4	0.5
		≥1	9.7 ± 1.9	11 ± 7.3	6.0	2.1	3.4	1.5
	250–400 GeV	≥0	4.15 ± 0.45	3.5 ± 1.5	1.3	0.5	0.5	0.2
		=0	1.70 ± 0.22	1.31 ± 0.72	0.65	0.16	0.25	0.10
		≥1	2.45 ± 0.45	2.6 ± 2.1	1.9	0.4	0.7	0.3
	400–600 GeV	≥0	0.62 ± 0.05	0.60 ± 0.40	0.37	0.07	0.12	0.08
	> 600 GeV	≥0	0.11 ± 0.01	-0.10 ± 0.12	0.12	0.01	0.03	0.01

Truth tagging with GNN



19

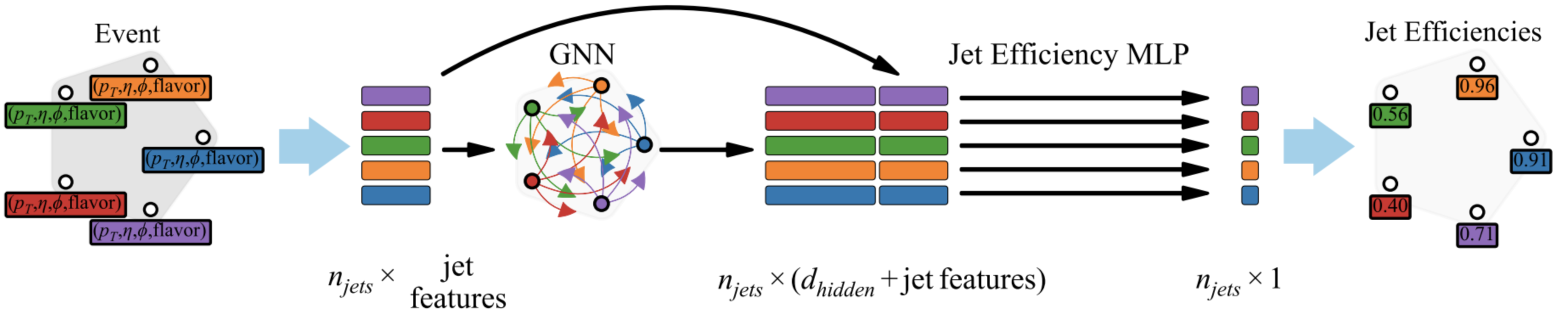
USING GNN'S FOR PARAMETRIZING ϵ_{jet}

Graph Neural Networks (GNN)

***Edges** set as ΔR*

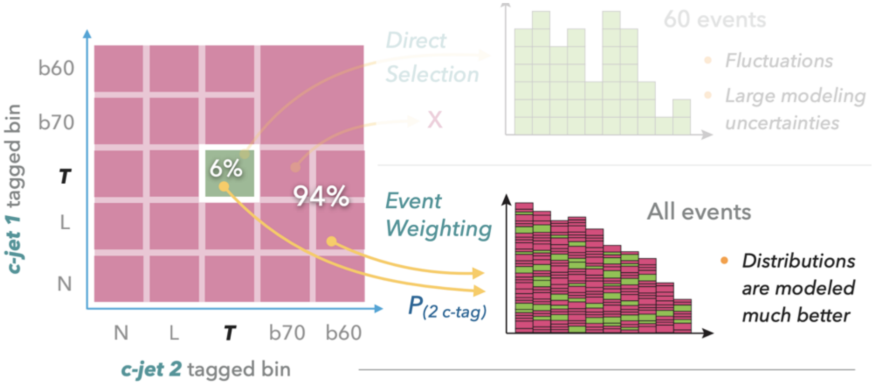
- *Each **node** is set as a jet*
- *Parameters set as input **features***

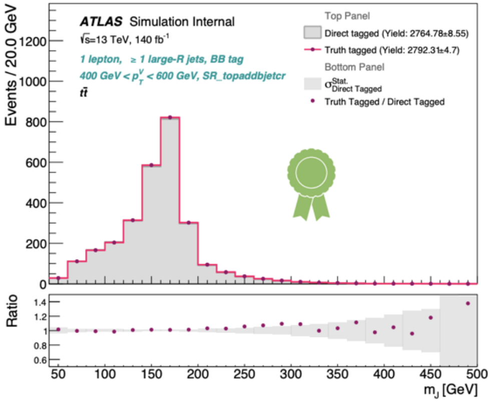
Reference: GNN Truth Tagging Paper

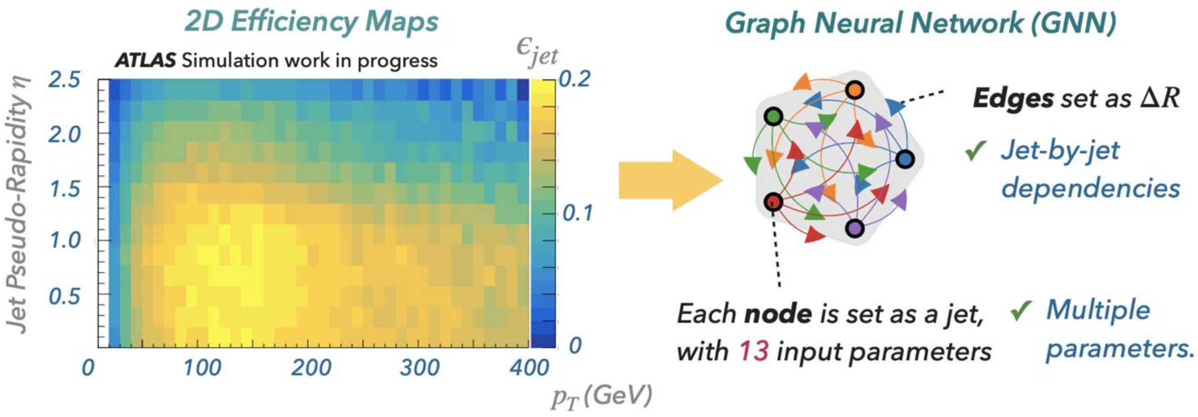


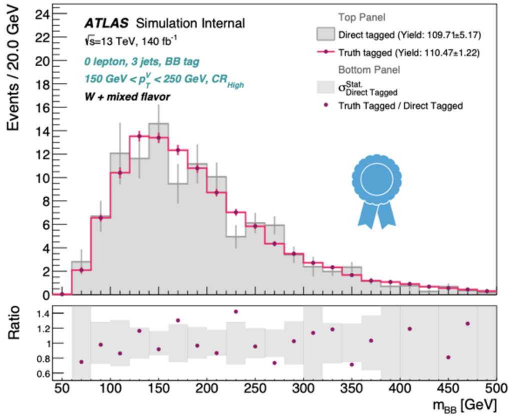
The key elements for the improvements – ML techniques

- Improving the Monte Carlo (MC) statistical uncertainties
  - Instead of rejecting the events failed the tagging requirement, calculate the probabilities for each jet to fall into a given tagging category
  - Previous iteration, use the 2D efficiency map, considered only pt and eta
  - Now, dedicated training with GNN with multiple paremeters
    - (much) improved closure with direct tagging
    - no need to consider the additional uncertainties to cover the unclosure









The results: Differential XSec (STXS) measurement

- Differential measurements of the VHbb cross-section in kinematic fiducial volumes defined in the simplified template cross-section (STXS) framework
- Reco pTV and nJet bin defined inline with the STXS binining
  - Good correspondence between truth and reco. bins (signal fraction in corresponding bins: 70-97%)
  - Strong reduction in correlations between STXS signal strengths
    - harmonized truth and recro level pT cuts
    - migrating the W(taunu)H event from 0L to 1L