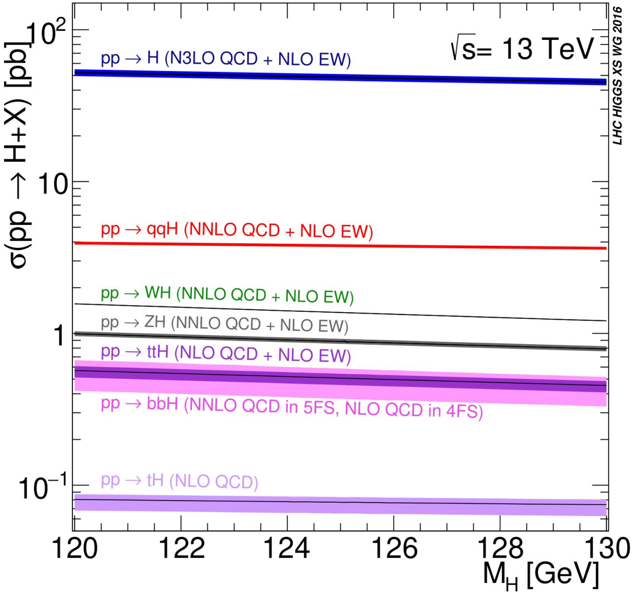
Search for H→bb decays in the VBF production mode

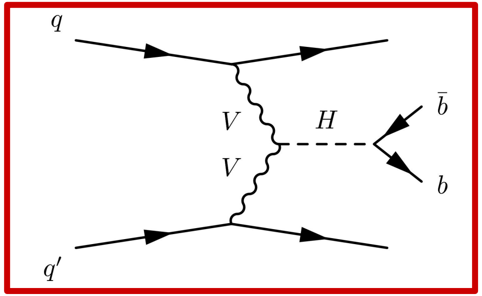
Presented by Matt Klein (University of Michigan)

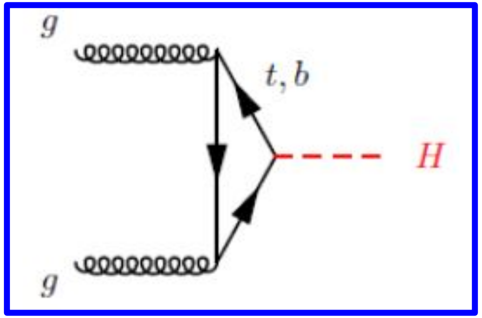
2021 July 12

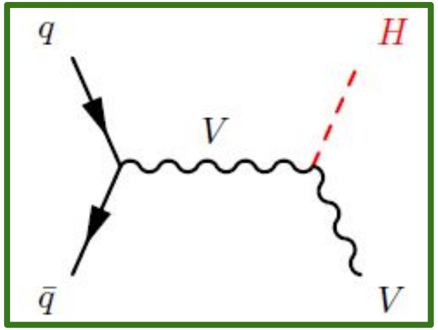
Introduction

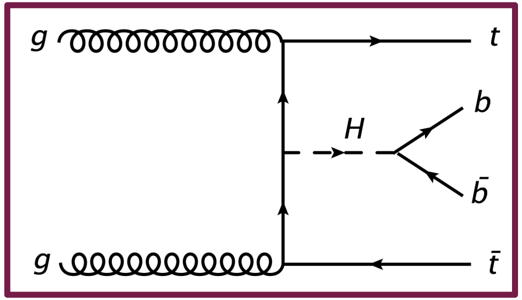
- H→bb observed by ATLAS and CMS in 2018, mostly with V(lep)H(bb)
- Other production modes allow for increased precision of Higgs coupling measurements, Higgs measurements at high p_T, and H→bb measurement not explicitly correlated with VH







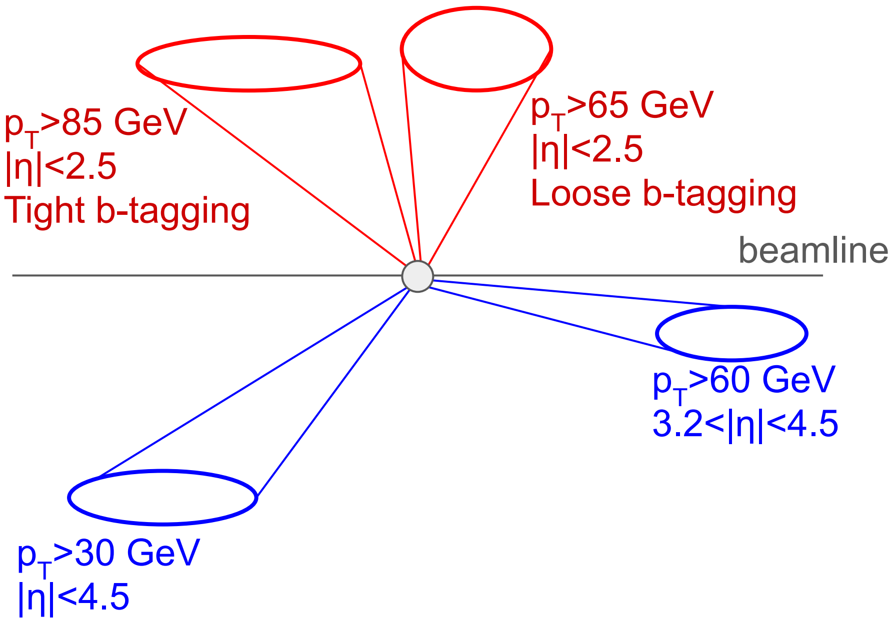




Analysis Channels and Selection

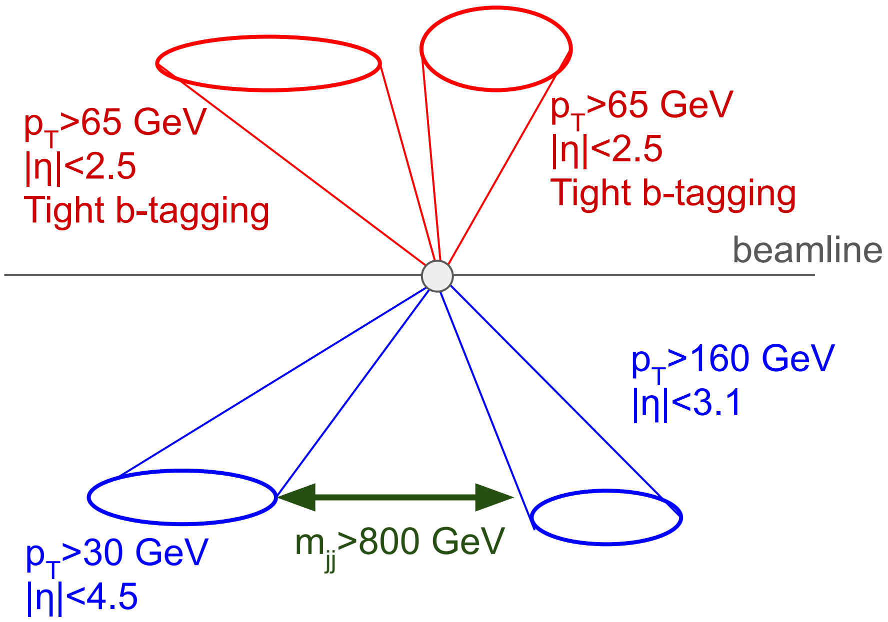
Forward Channel

Events contain a high p_T forward jet



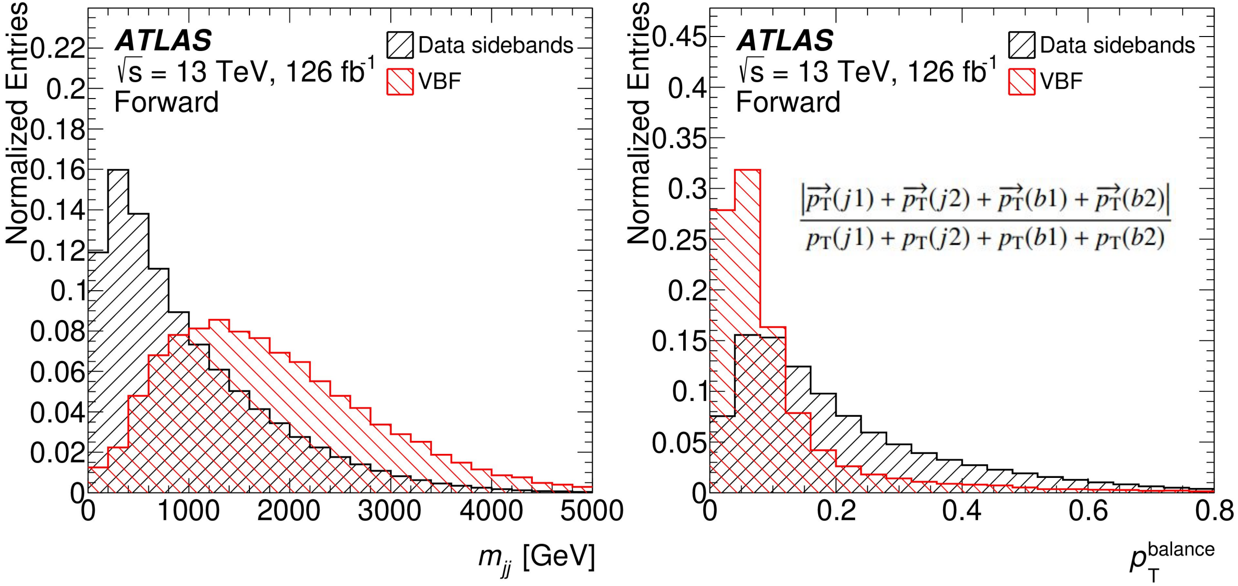
Central Channel

Events do not contain a high p_T forward jet



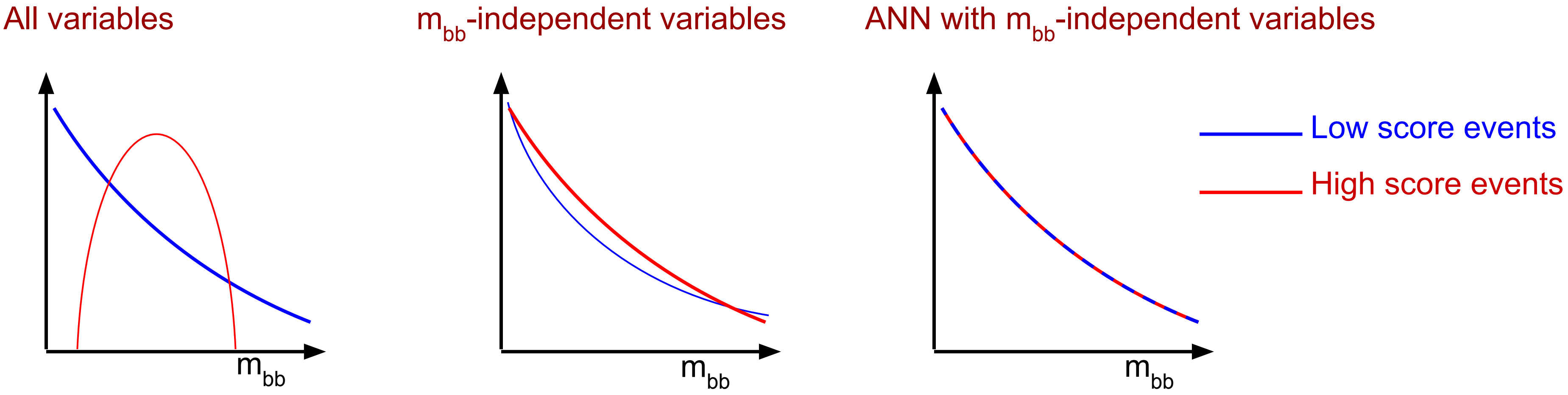
Signal/Background Discrimination

- MVA to categorize events, based on set of input kinematic variables
- Strongest discriminants based on angular separation of jets or on the presence of extra jets or soft emissions
- Extra discrimination comes from, for example, quark/gluon tagging



Adversarial Neural Network

- MVA variables chosen to be largely uncorrelated with m_{bb}
- Residual correlations still possible, particularly when combining variables in MVA
- Train an **adversarial neural network** in each channel, effectively adding a term to the loss function to explicitly penalize a network that correlates m_{bb} and score
- Training performed between signal and data sidebands

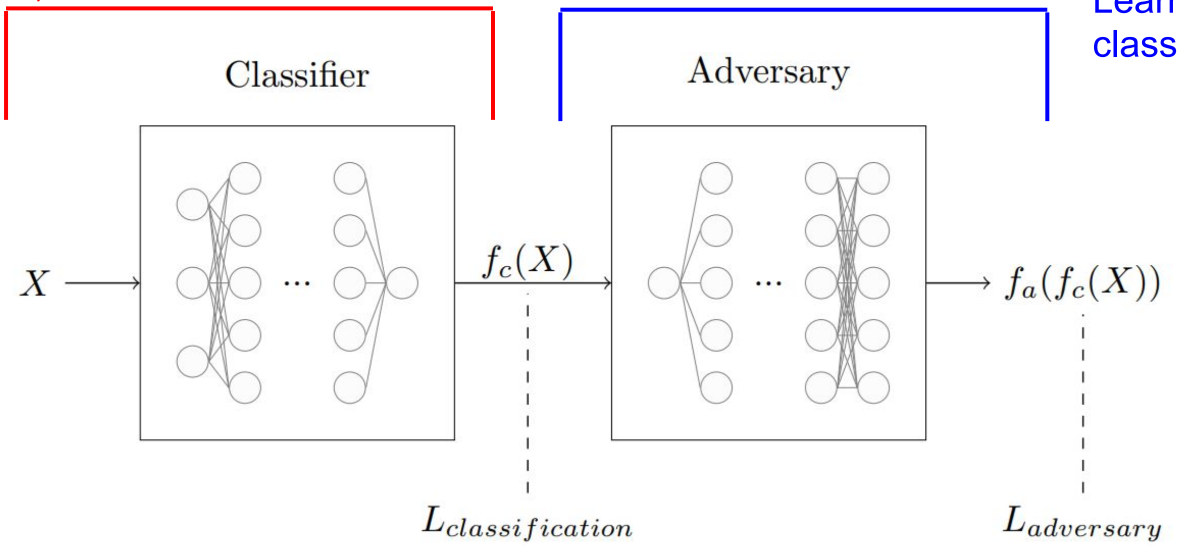


ANN Architecture

$L = L_{classifier} - \lambda L_{adversary}$

VBF H vs. background,
with 12 inputs

Learn mass from
classifier response



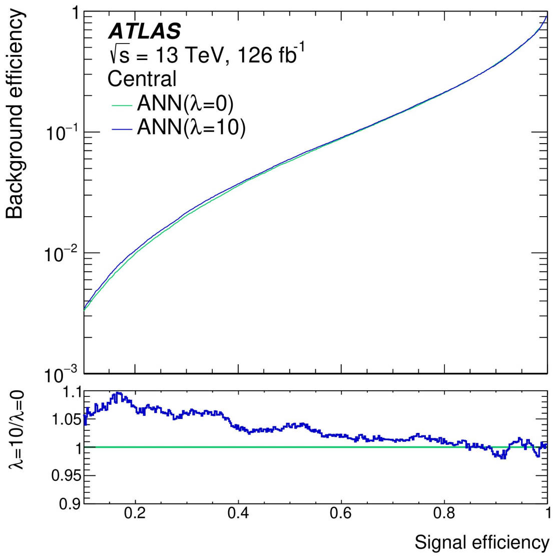
Input → BatchNorm → Dense (256) → Sigmoid → Dense (128) → Softmax (10)

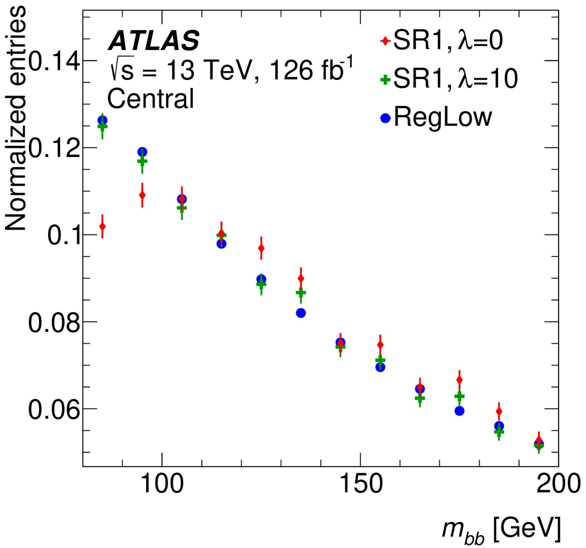
[Phys. Rev. D 96, 074034 (2017)]

ANN Performance

ROC curves: Performance using **adversarial** or **non-adversarial** training

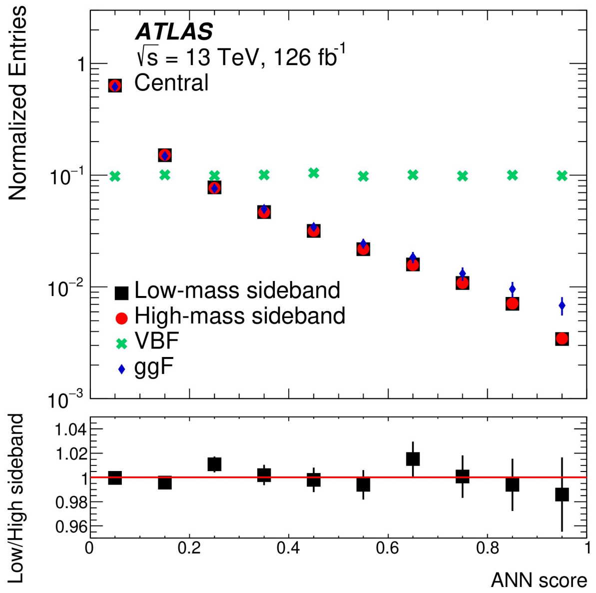
M_{bb}: Shape of **low-score region**, non-**adversarial high-score region**, and **adversarial high-score region**





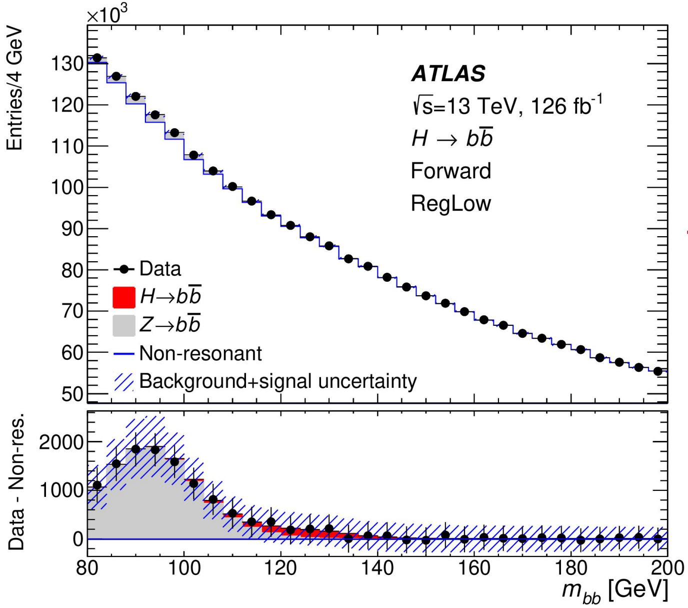
Signal/Background Discrimination

- Consistent shapes in high and **low** mass sidebands
- ggF similar to multijet background
- VBF flat by design

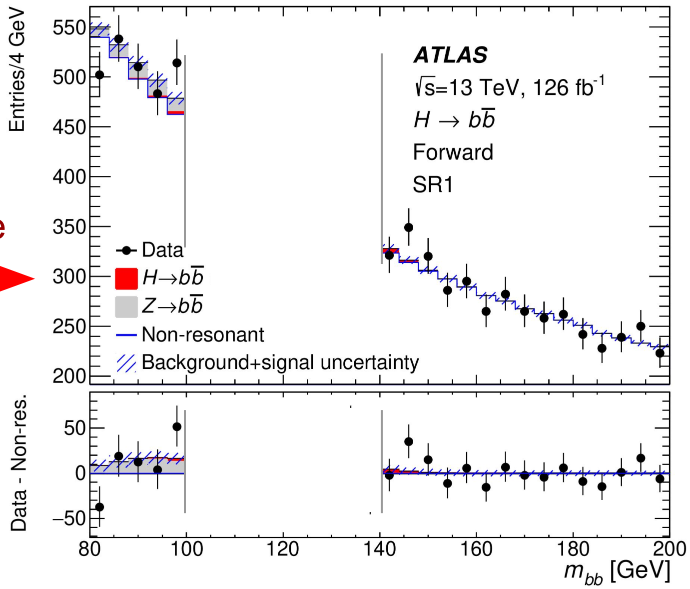


Fit Strategy

- Fit 4 high score regions + one low score region in each channel (10 regions total)
- Score and m_{bb} uncorrelated → constrain background template in high score signal regions using high-statistic low score control region



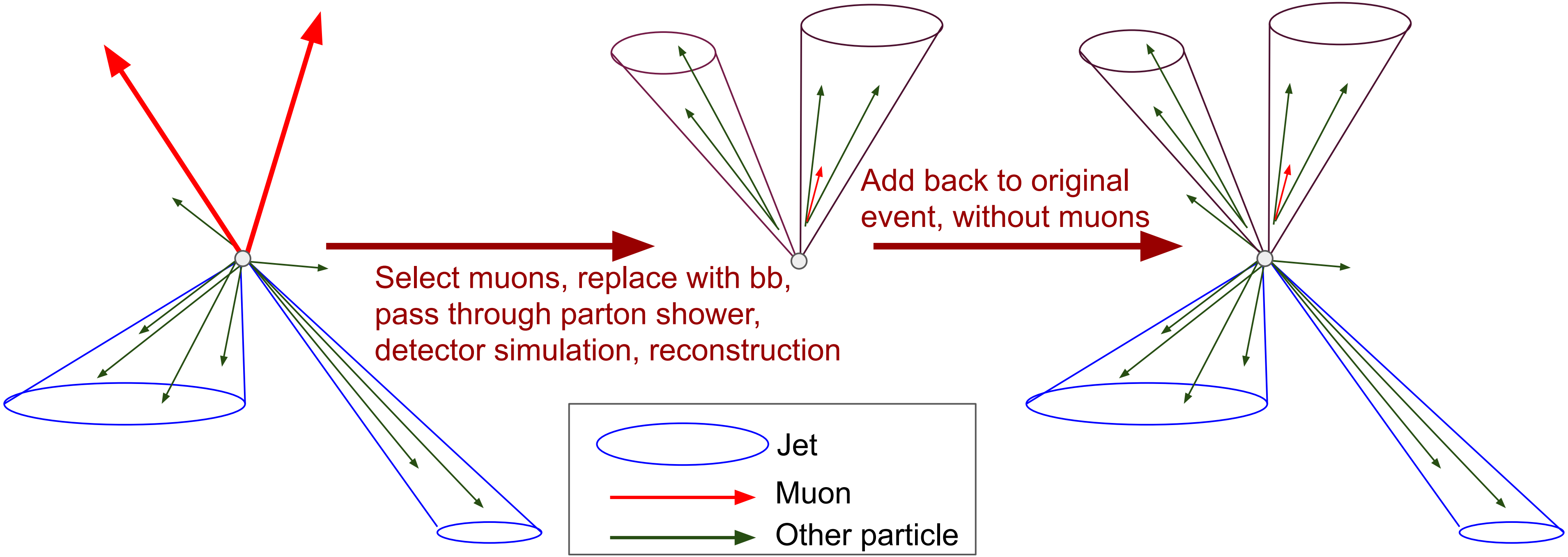
Take background shape



Z→bb Background

- Potentially significant mismodeling and systematic uncertainties
- Due to trigger limitations, cannot constrain Z in fits

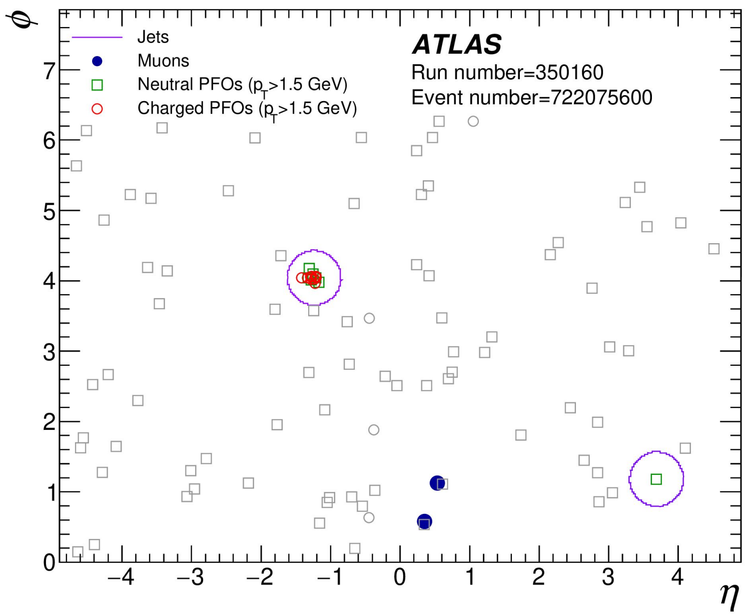
  →Data-driven approach: estimate Z→bb from Z→μμ (embedding)



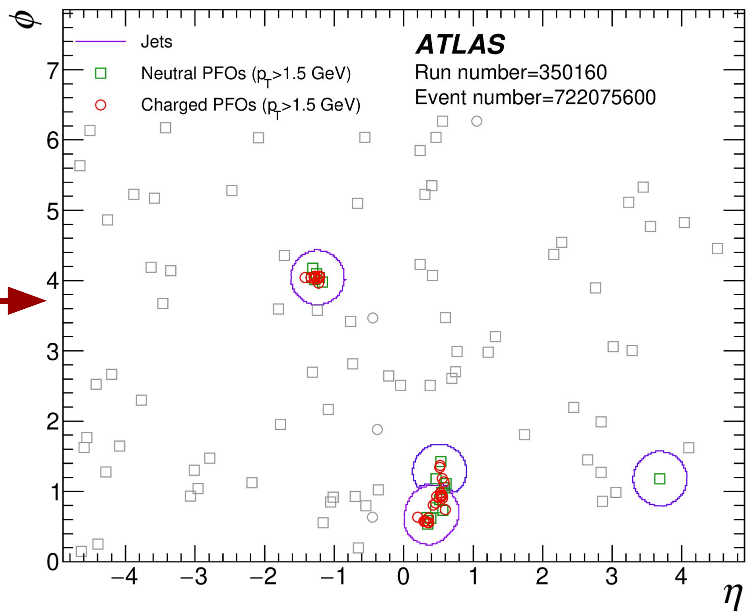
Z→bb Background Example

Example event in data

Before

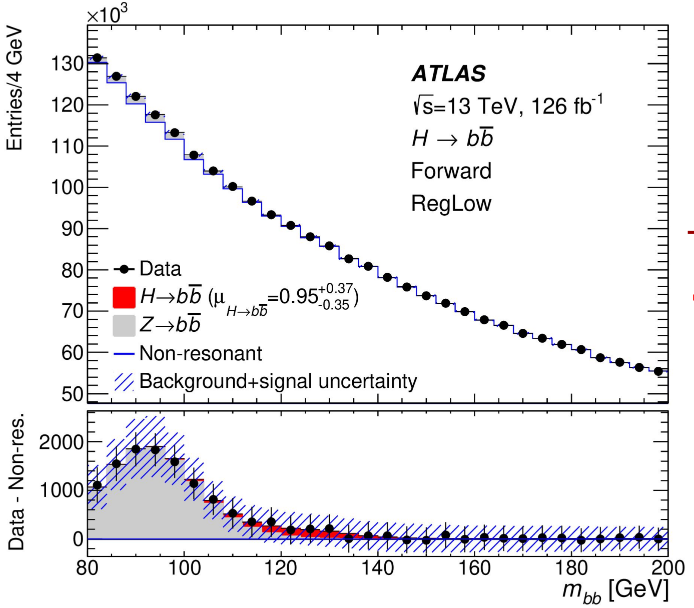


After

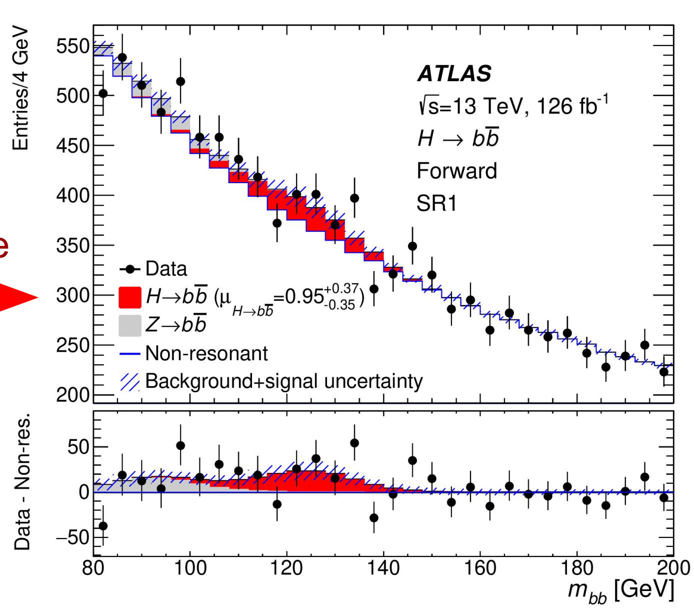


Fits (Forward Channel)

- Fit 4 high score regions + one low score region in each channel (10 regions total)
- Score and m_{bb} uncorrelated → constrain background template in high score signal regions using high-statistic low score control region

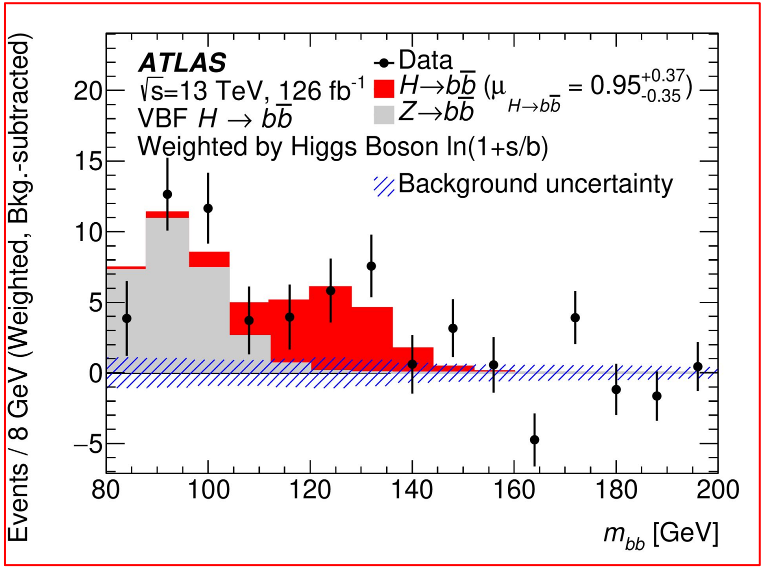


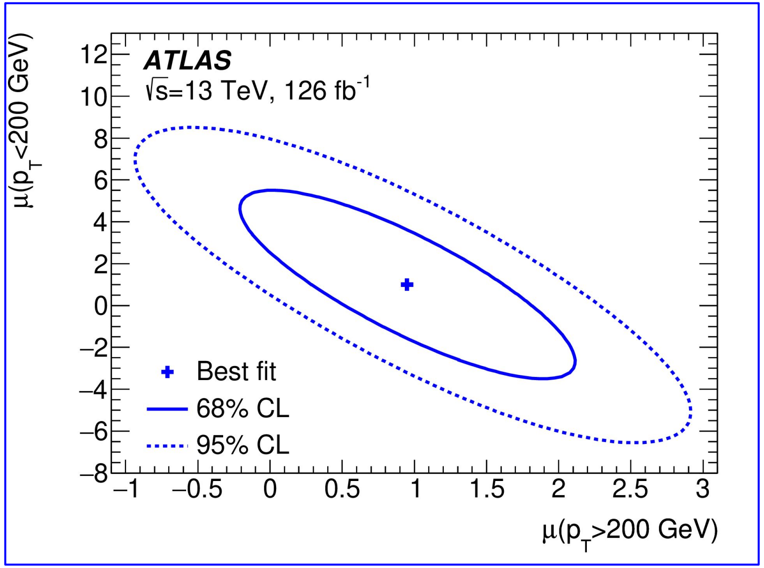
Take background shape →



Results

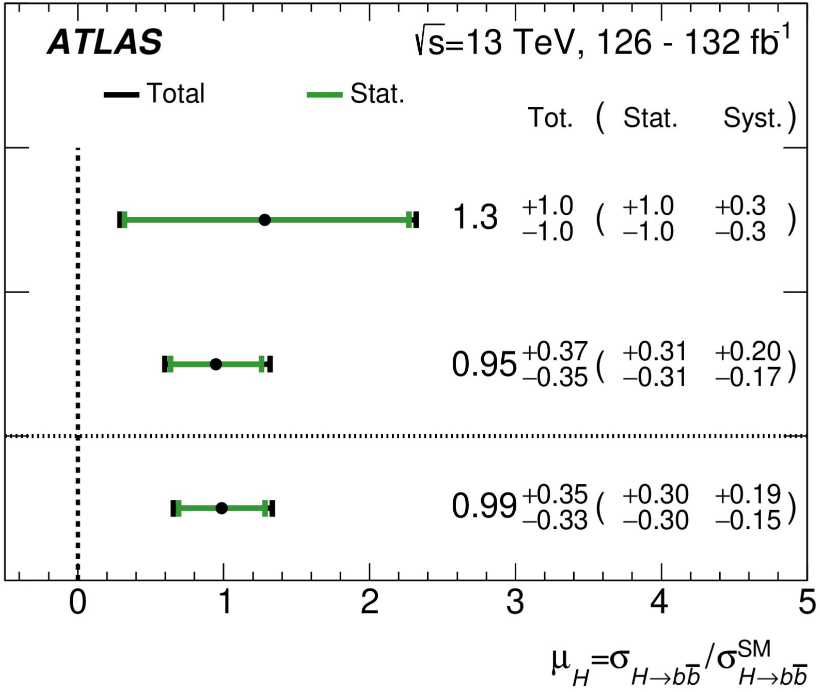
Results	Inclusive Production	VBF Production	$p_T > 200$ GeV
Expected significance	2.9σ	2.8σ	2.3σ
Observed significance	2.7σ	2.6σ	2.2σ
Expected signal strength	$1.00^{+0.37}_{-0.36}$	$1.00^{+0.38}_{-0.37}$	$1.00^{+0.45}_{-0.43}$
Observed signal strength	$0.95^{+0.37}_{-0.35}$	$0.95^{+0.38}_{-0.36}$	$0.93^{+0.45}_{-0.43}$





Combination

- Combination of all-hadronic VBF H→bb and VBF H+γ results
- Statistical uncertainties dominant
- Observe (expect) 3.0σ (3.0σ) for H→bb production and 2.9σ (2.9σ) for VBF H→bb

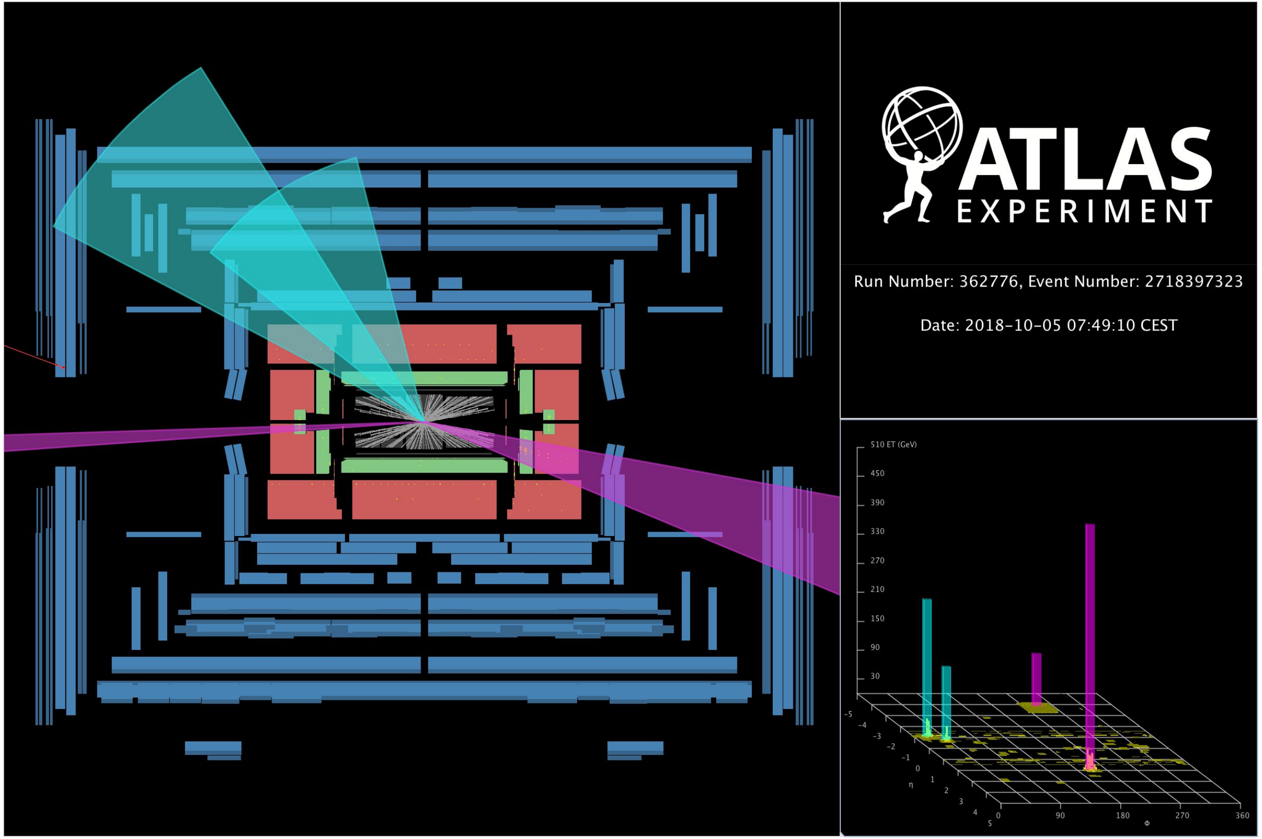


Conclusion

Many new Higgs results from ATLAS in 2020 for bb final states

For VBF+ggF, measured 3σ for H→bb production

See paper for more information



Backup

Citations

ATLAS Collaboration, Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at √s=13 TeV. *Eur. Phys. J. C* 81, 537 (2021). https://doi.org/10.1140/epjc/s10052-021-09192-8

The ATLAS collaboration, Search for Higgs boson production in association with a high-energy photon via vector-boson fusion with decay into bottom quark pairs at √s = 13 TeV with the ATLAS detector. *J. High Energ. Phys.* 2021, 268 (2021). https://doi.org/10.1007/JHEP03(2021)268

Gabrielli, E., Mele, B., Piccinini, F. *et al.* Asking for an extra photon in Higgs production at the LHC and beyond. *J. High Energ. Phys.* 2016, 3 (2016)

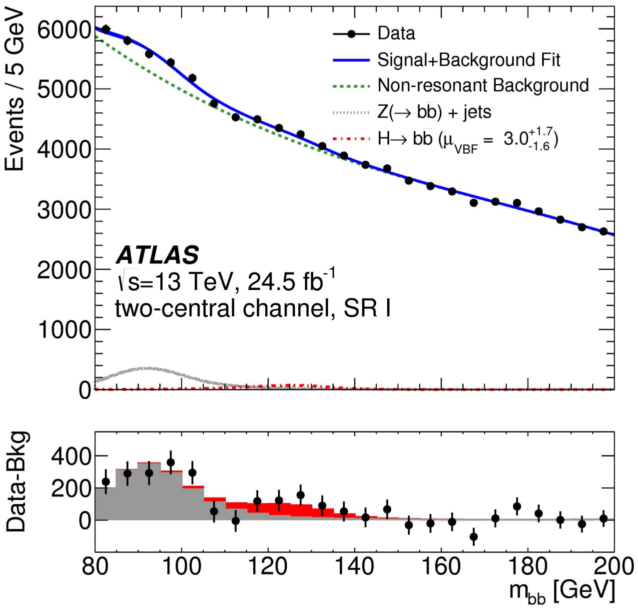
A. D. Bukin, Fitting function for asymmetric peaks, 2007, arXiv: 0711.4449 [physics.data-an]

C. Shimmin et al., Decorrelated jet substructure tagging using adversarial neural networks, Phys. Rev. D 96 (2017), arXiv: 1703.03507

arXiv:1807.08639

History: VBF Hbb and VBF Hbb+γ

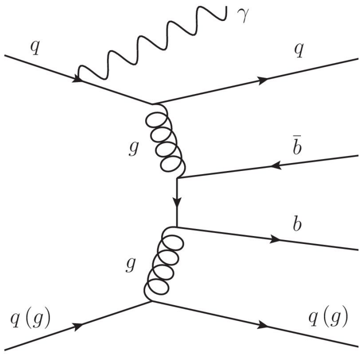
Large, difficult to model multijet background
Smaller, peaked Z→bb background near signal
Trigger on jets

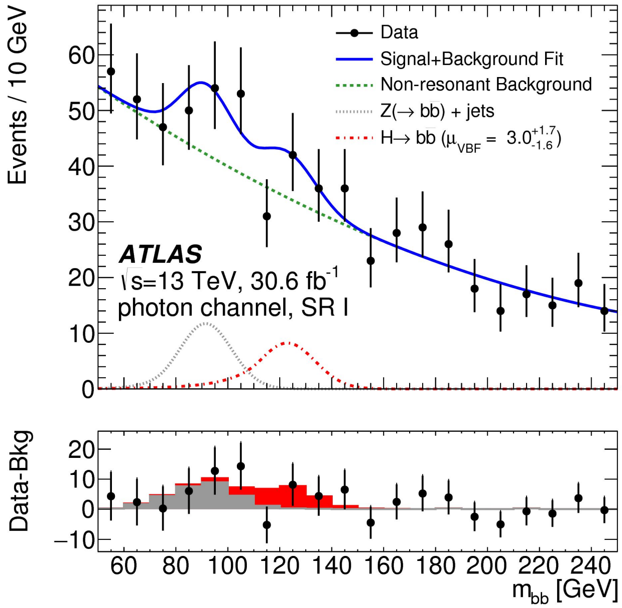


Larger Z→bb background

Mainly γ+jets background - gg-initiated diagrams suppressed

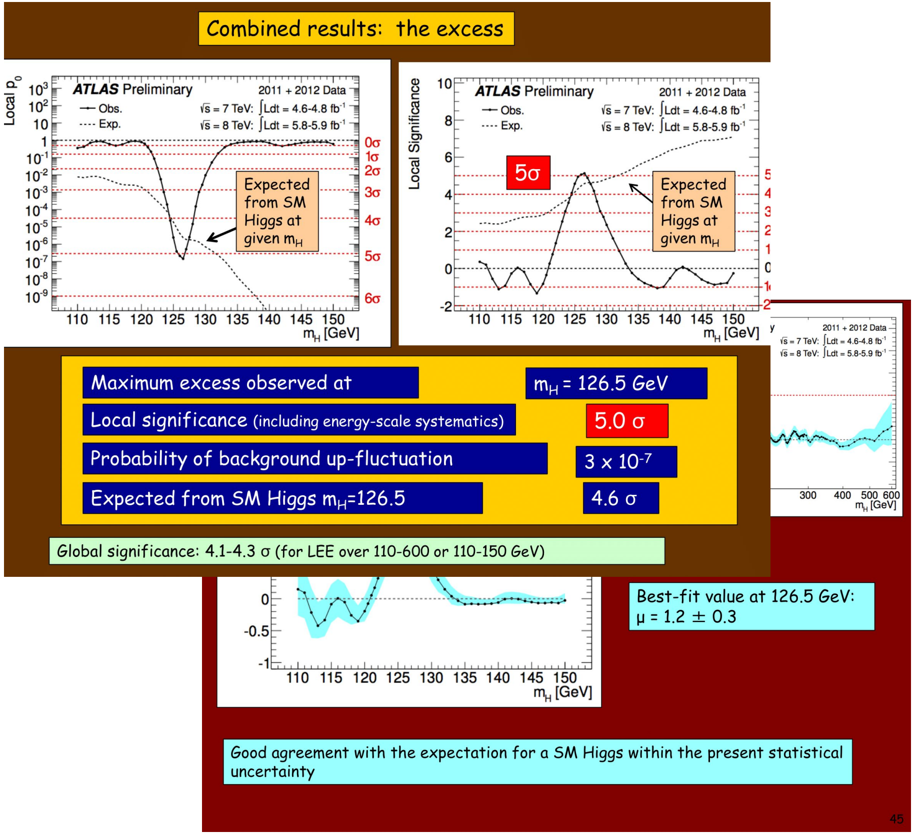
Trigger on γ+jets





ATLAS and the Higgs boson

2011-2012 data



Now

Combined Measurements of Higgs Boson production and decays

Channels included in the combination:

Analysis	Dataset	$\mathcal{L}$ [fb^{-1}]
$H \to \gamma\gamma$	2015–2018	139
$H \to ZZ^* \to 4\ell$		
$VH, H \to b\bar{b}$		
$H \to \mu\mu$		
VBF, $H \to inv$		
$H \to WW^* \to e\nu\mu\nu$	2015–2016	36.1
$H \to \tau\tau$		36.1
VBF, $H \to b\bar{b}$		24.5 – 30.6
$t\bar{t}H, H \to b\bar{b}$ and $t\bar{t}H$ multilepton		36.1

(i) Global signal strength

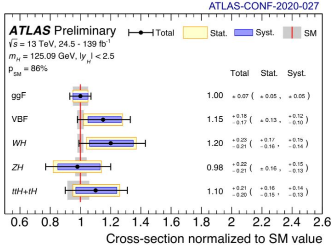
Describing a common scaling of the expected Higgs boson yields in all processes (σ x BR) (SM: μ = 1.0)

$\mu = 1.06 \pm 0.07$

[± 0.04 (stat) ± 0.03 (exp) $^{+0.05}_{-0.04}$ (sig.th) ± 0.02 (bkg.th)]

(ii) Production Cross Sections
(assume SM branching ratios)

ATLAS-CONF-2020-027



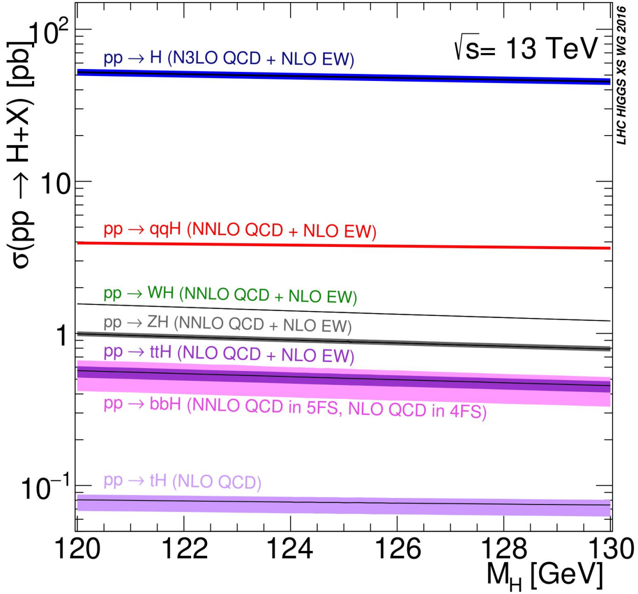
All major production processes observed
(significance > 5σ)

First observation of WH production: 6.3σ (5.2σ exp)

K. Jakobs, ATLAS Experiment, ICHEP 2020

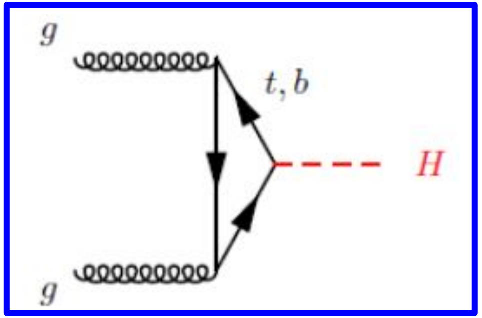
8

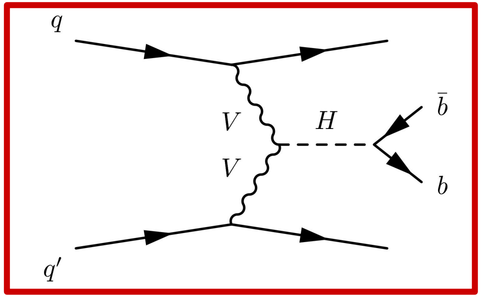
Higgs Production Modes

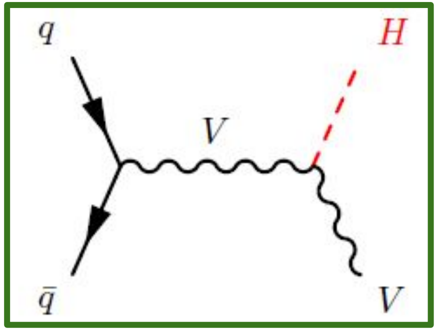


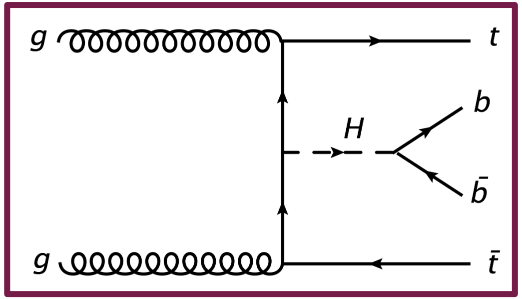
Expect H→bb 58% of time

Decreasing cross-section







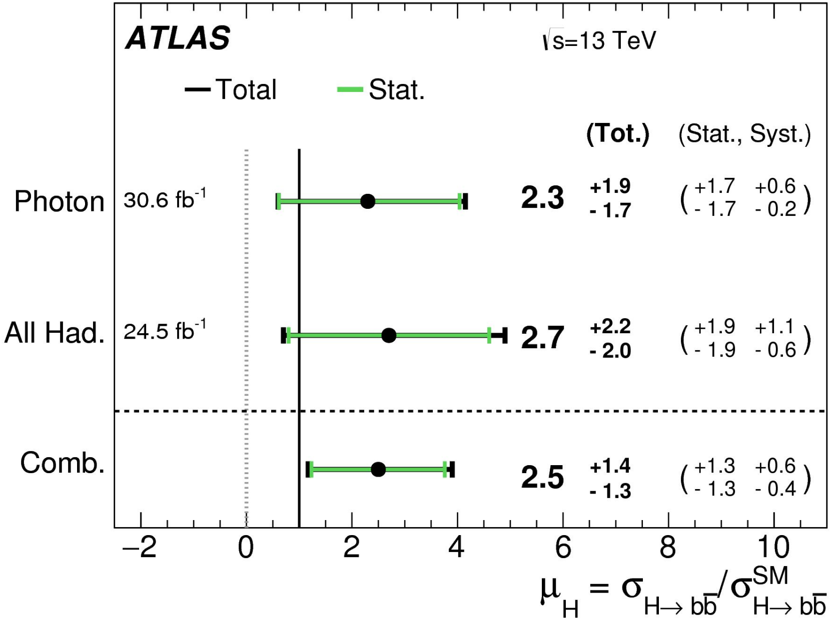


Decreasing background

[arXiv:1807.08639]

History: VBF H→bb

Results	Inclusive production		
	All-hadronic	*Photon*	Combined
Expected significance	0.5σ	0.6σ	0.8σ
Observed significance	1.4σ	1.3σ	1.9σ



ANN Architecture

- Pre-train the classifier (CL) with binary cross entropy loss (adversary parameters frozen).
- Pre-train adversary (AD) with categorical cross entropy loss (CL parameters frozen).
- Train CL and AD together with loss $L = L_{classifier} - \lambda L_{adversary}$ following these sub-steps for each of the 50 epochs:
  - Train CL with loss $L = L_{classifier} - \lambda L_{adversary}$, keeping AD weights frozen.
  - Train AD with loss $L = L_{adversary}$, , keeping CL weights frozen.

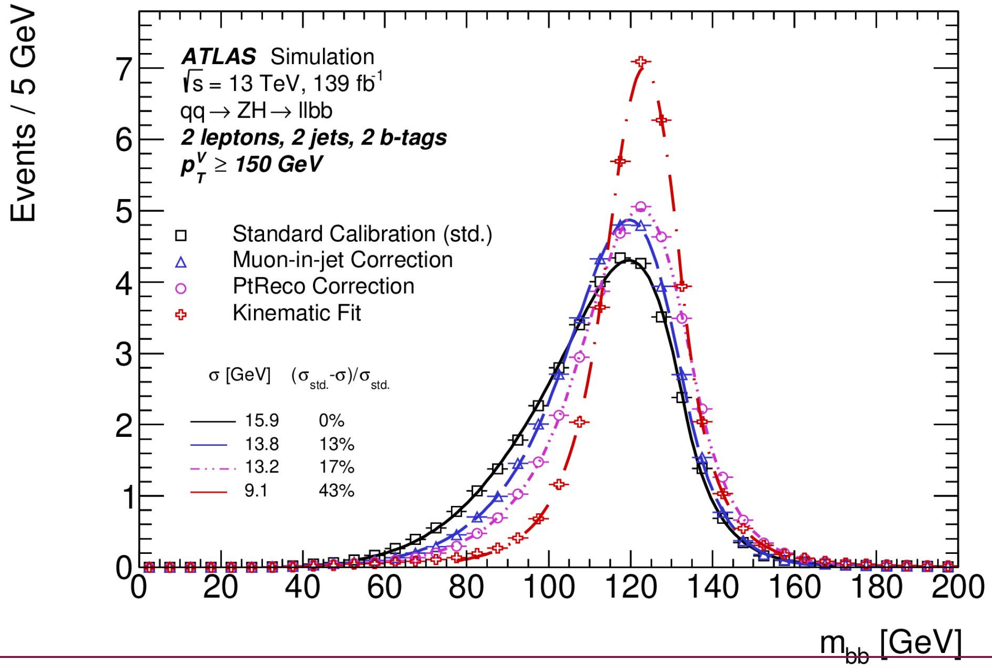
Input → BatchNorm → Dense (256) → Sigmoid → Dense (128) → Softmax (10)

The optimizer is Adam. lr=0.001, 821 beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.0, amsgrad=False

B-jet Corrections

m_{bb} resolution using different b-jet corrections

VBF and VBF+ɣ use muon-in-jet and PtReco



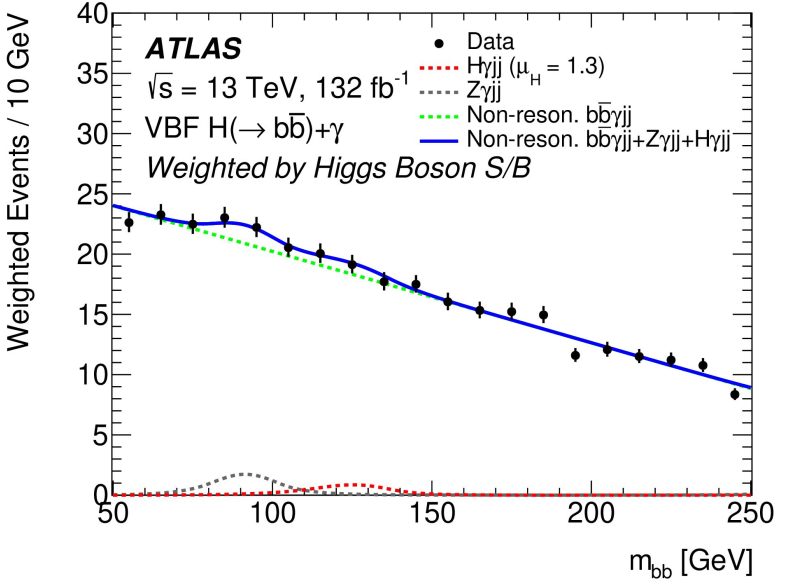
arXiv:2007.02873

VBF+γ: Selection and results

Left: Full event preselection

Right: weighted sum of signal regions, including non-resonant background

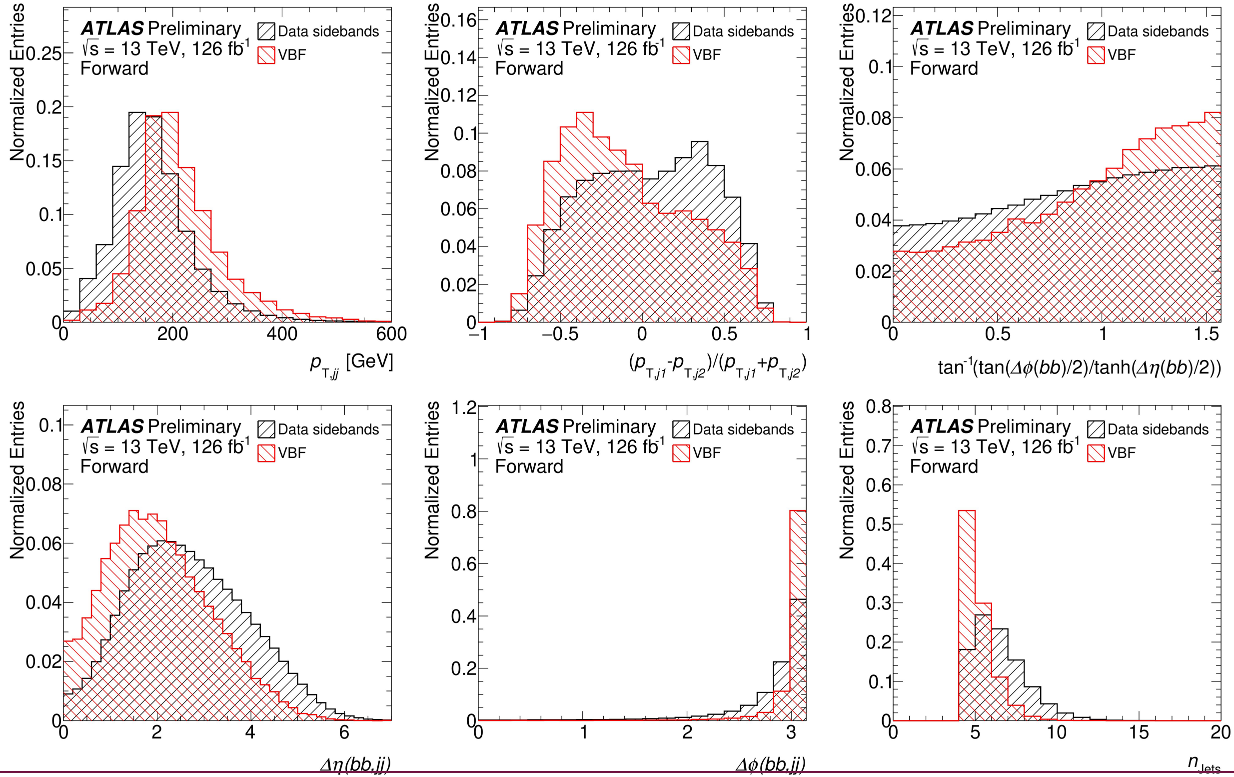
Trigger	L1	≥ 1 photon with $E_T > 22$ GeV
	HLT	≥ 1 photon with $E_T > 25$ GeV ≥ 4 jets (or ≥ 3 jets and ≥ 1 b-jet) with $E_T > 35$ GeV and $\lvert\eta\rvert < 4.9$ $m_{jj} > 700$ GeV
Offline		≥ 1 photon with $E_T > 30$ GeV and $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$ ≥ 2 b-jets with $p_T > 40$ GeV and $\lvert\eta\rvert < 2.5$ ≥ 2 jets with $p_T > 40$ GeV and $\lvert\eta\rvert < 4.5$ $m_{jj} > 800$ GeV $p_T(b\bar{b}) > 60$ GeV No electrons ($p_T > 25$ GeV, $\lvert\eta\rvert < 2.47$) or muons ($p_T > 25$ GeV, $\lvert\eta\rvert < 2.5$)



All-hadronic VBF H→bb Variable List

11 (12) input variables for Forward (Central) channel. In the Forward channel, N_{trk}^{j1} is not included, as one of the jets is always in the forward region and has no associated tracks

- m_{jj}
- $p_{T,jj}$
- p_T balance
- $\Delta\eta(bb,jj)$
- $\Delta\varphi(bb,jj)$
- N jets
- $(p_T^{j1}-p_T^{j2})/(p_T^{j1}+p_T^{j2})$
- $\tan^{-1}(\tan(½\Delta\varphi^{bb})/\tanh(½\Delta\eta^{bb}))$
- $N_{trk}^{j1\,(2)}$ (q/g tagging)
- $\min(\Delta R(j^{1\,(2)}$, extra jet))

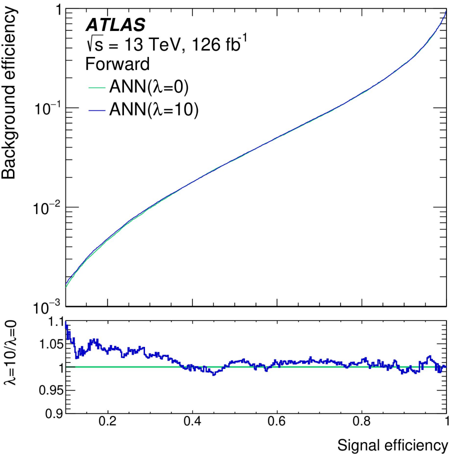


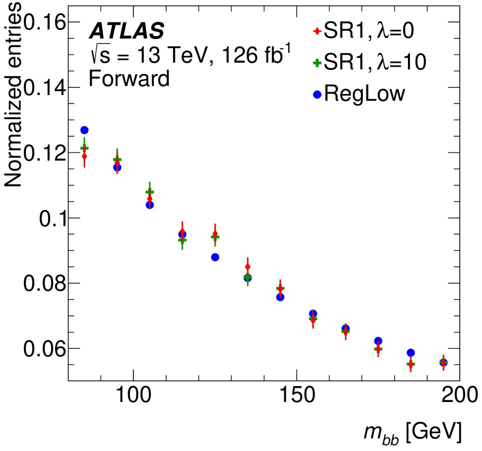
ANN Performance

ROC curves: Performance using **adversarial** or **non-adversarial** training

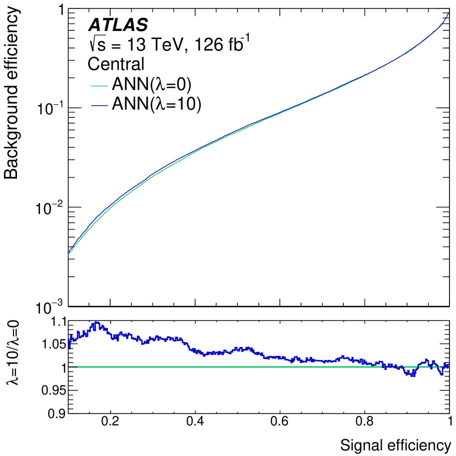
M_{bb}: Shape of **low-score region**, non-**adversarial high-score region**, and **adversarial high-score region**

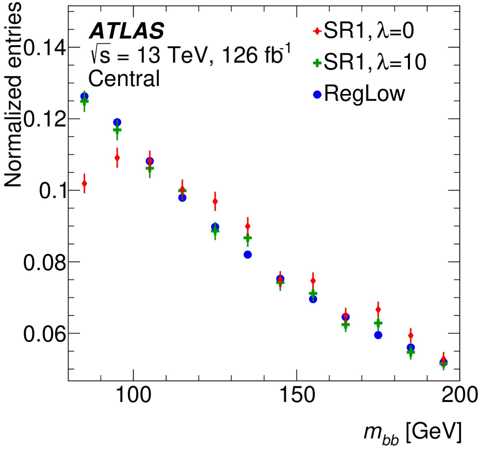
Forward





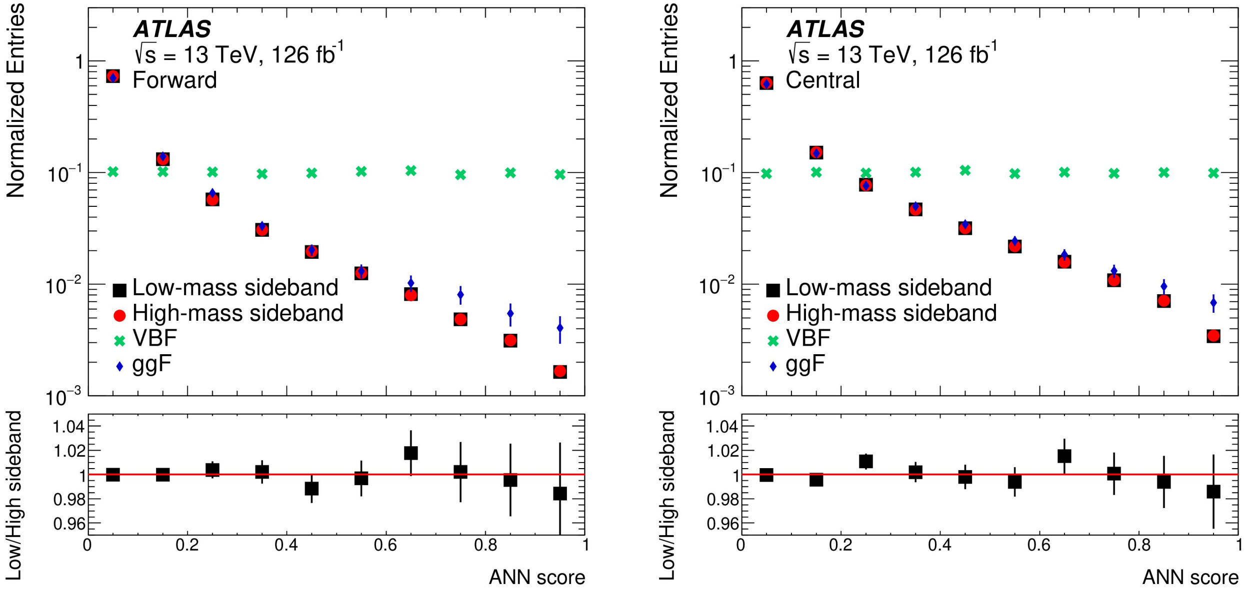
Central





Signal/Background Discrimination

- Consistent shapes in high and **low** mass sidebands
- ggF similar to multijet background
- VBF flat by design



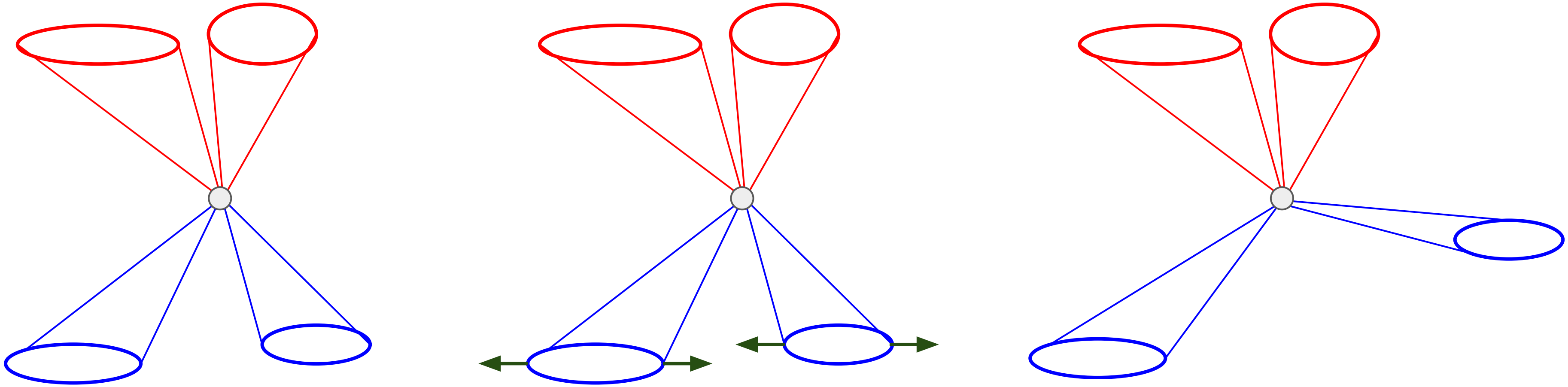
Background Uncertainties

Define CRs to measure any potential bias induced by excluding Higgs mass window

Select events with low m_{jj} and no forward jets

Smear VBF jet η to mimic SR kinematics

Reweight event-kinematics to remove residual SR-CR differences

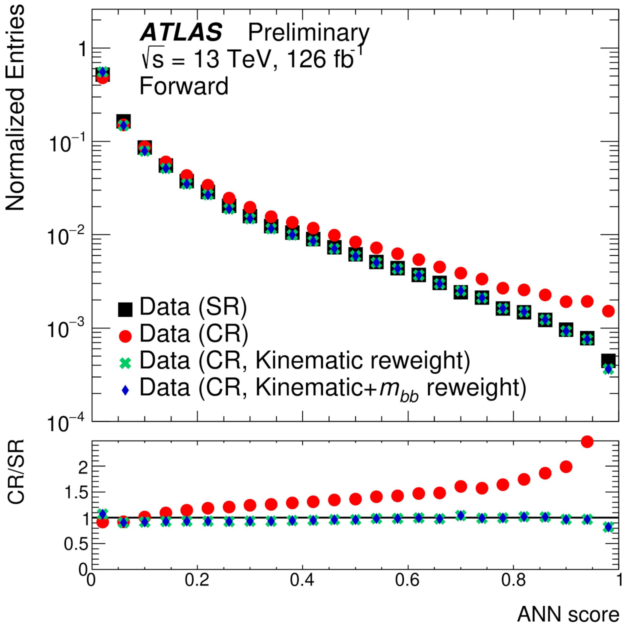


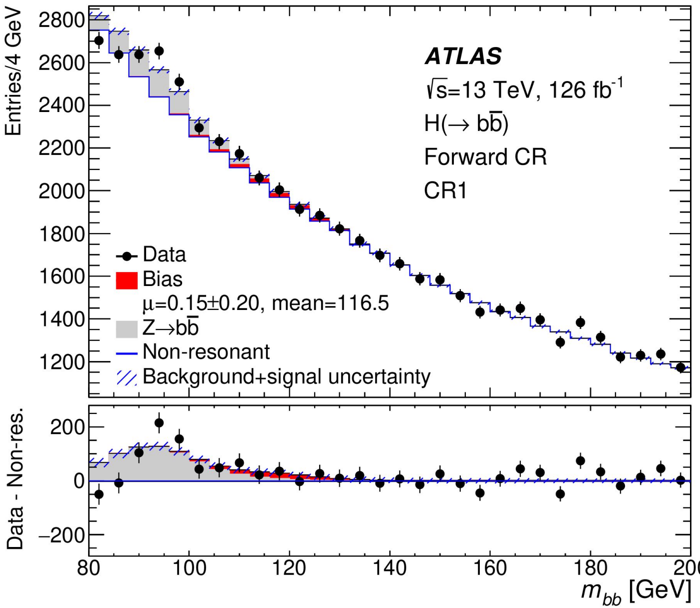
Fit signal+background to CRs to measure maximum possible bias

Classifier and CRs

Want CRs and SRs to be kinematically consistent

Score distribution before kinematic CR→SR reweighting, after reweighting (except m_{bb}), and after reweighting m_{bb}

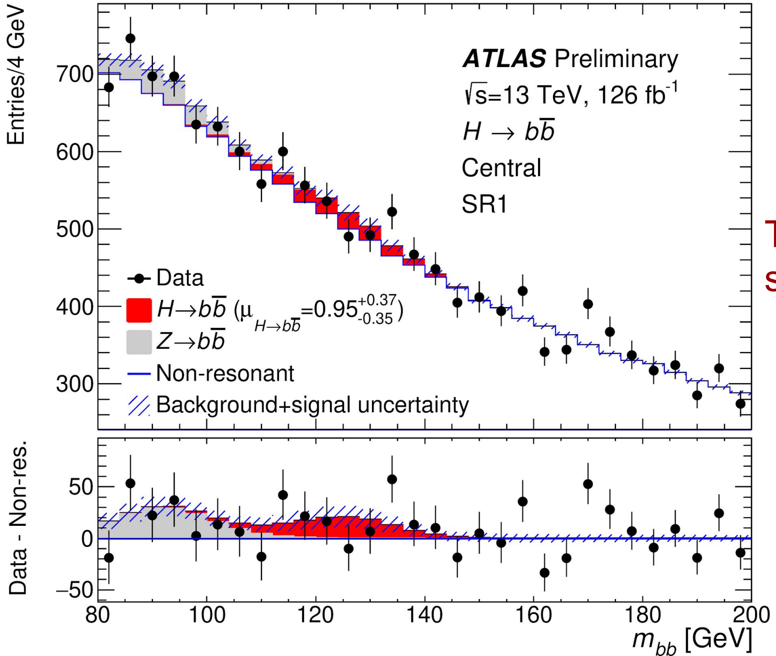




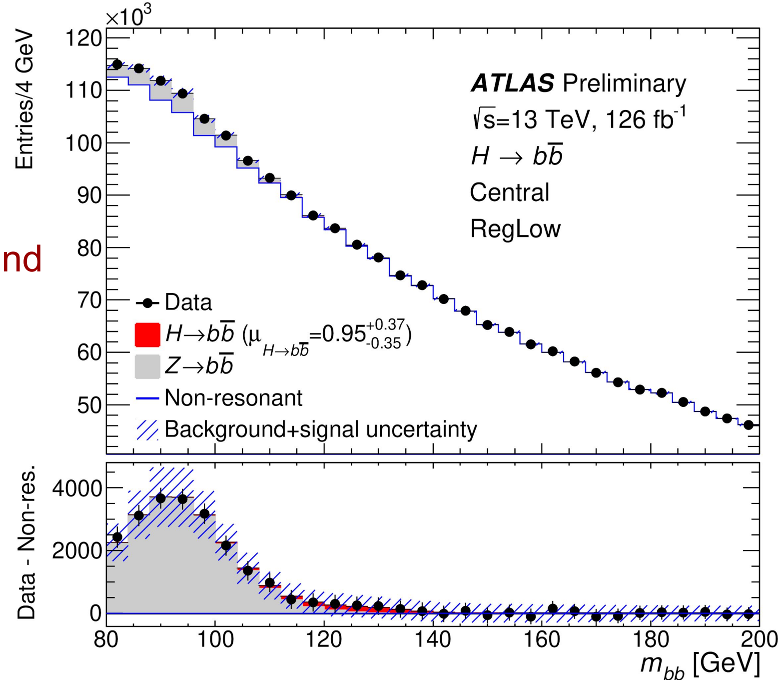
Fits (Central Channel)

- Score and m_{bb} uncorrelated → take background template in high score signal regions from high-statistic low score control region

- Fit 4 high score regions + one low score region in each channel (10 regions total)



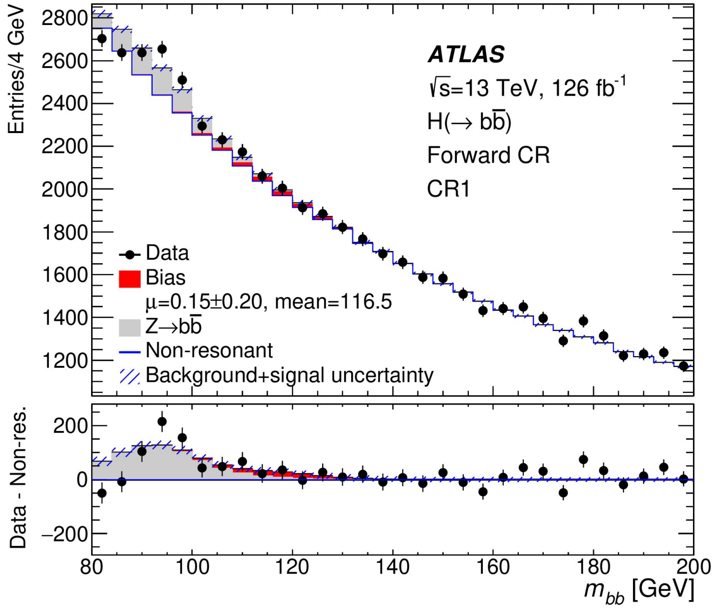
Take background shape from

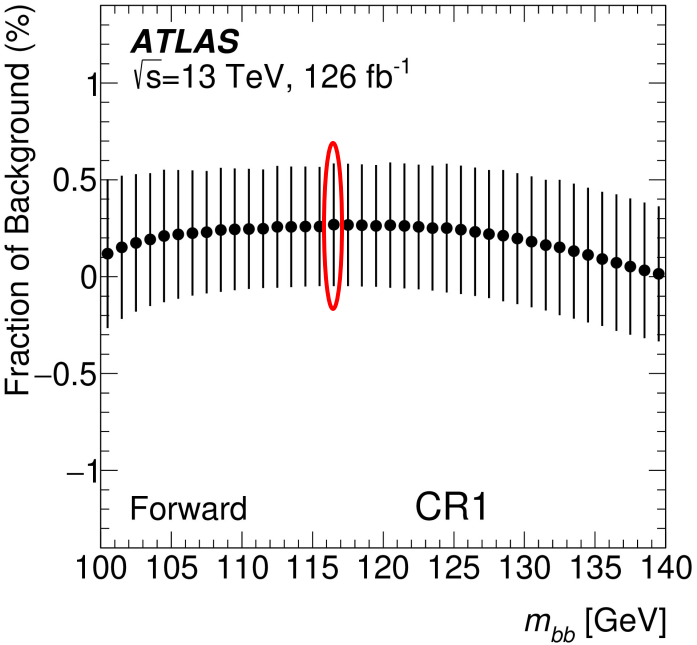


Classifier and CRs

Fit signal+background to CRs to measure potential bias

Scan over $100<m_{bb}<140$ GeV and take max extracted signal as uncertainty in SR

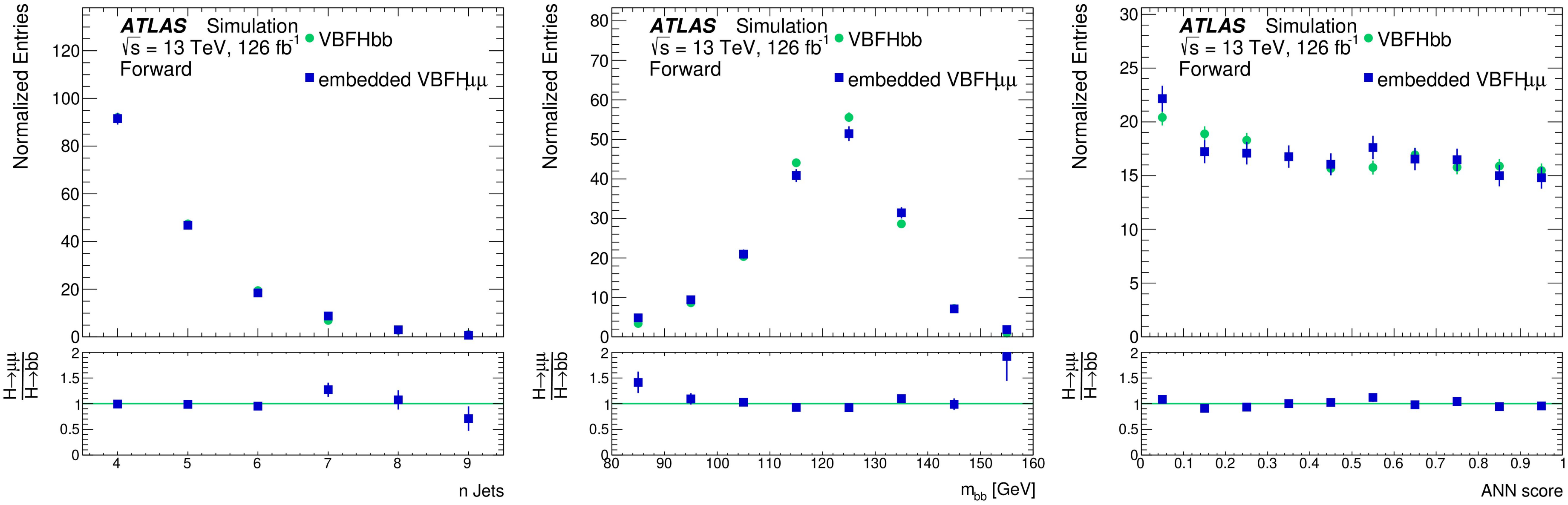




Z→bb Closure

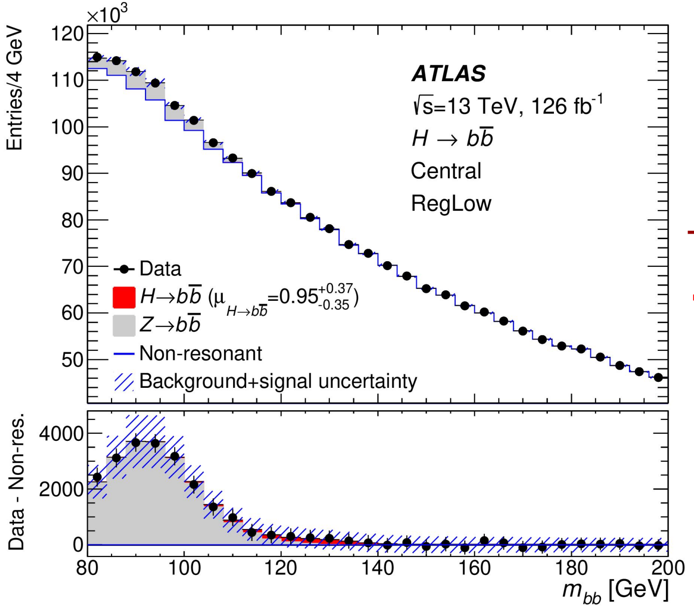
Closure checked in both **Zμμ** vs. **Zbb** and **Hμμ** vs. **Hbb**

Latter gives higher statistics in VBF-like topology

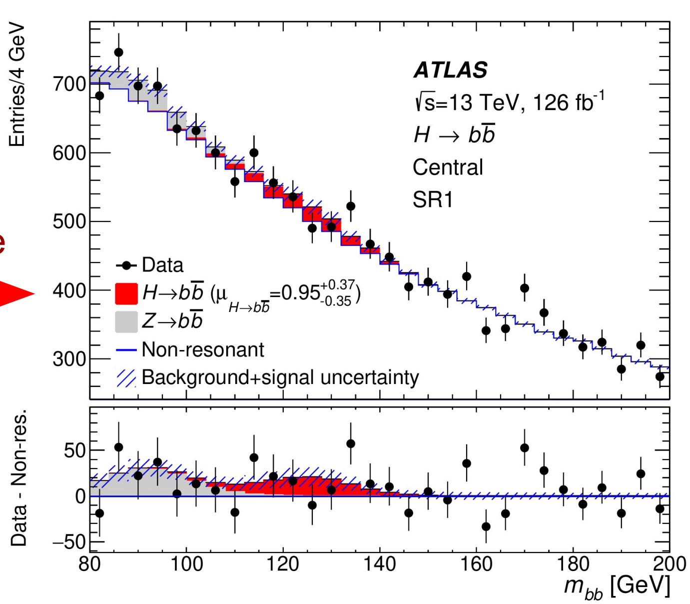


Fits (Central Channel)

- Fit 4 high score regions + one low score region in each channel (10 regions total)
- Score and m_{bb} uncorrelated → constrain background template in high score signal regions using high-statistic low score control region



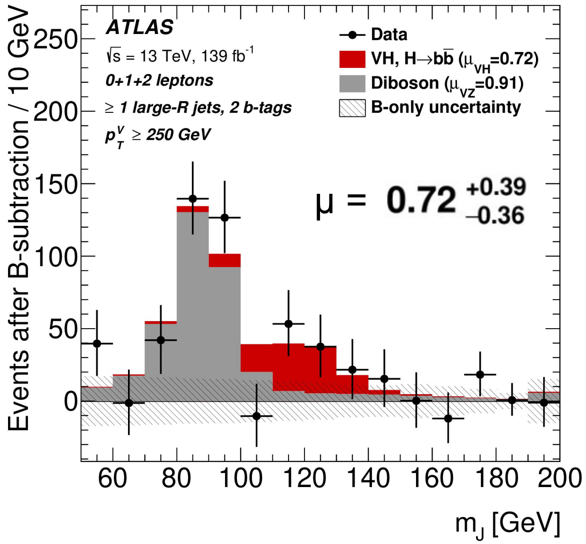
Take background shape



Other H→bb Results

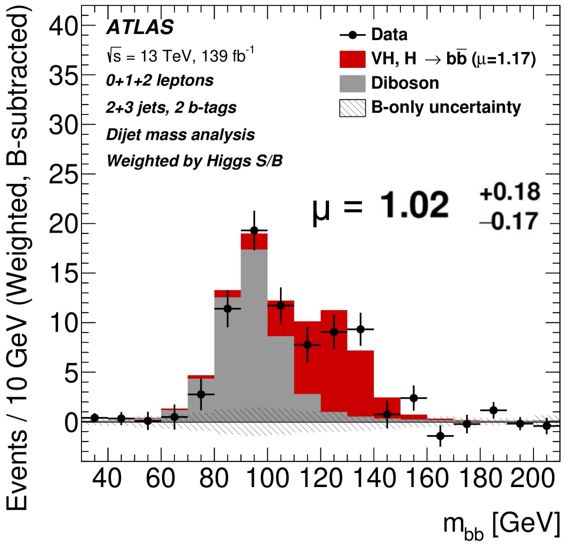
V(lep)H(bb) and ttH(bb) have also released results with the full Run 2 dataset

Boosted V(lep)H(bb)
(p_T^H>250 GeV)



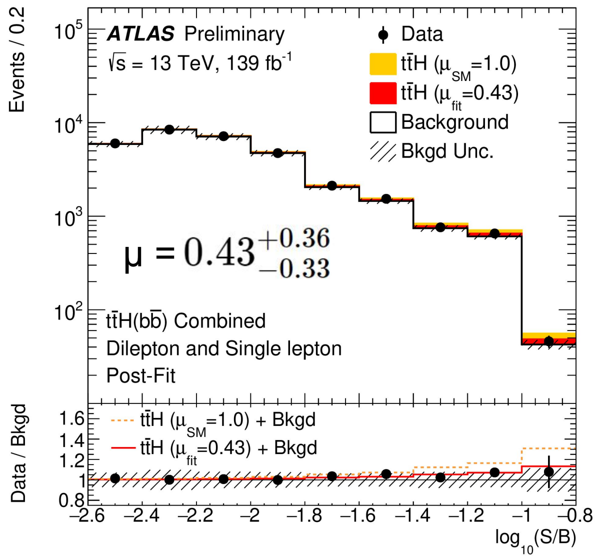
arXiv:2008.02508

Resolved V(lep)H(bb)



arXiv:2007.02873

ttH(bb)

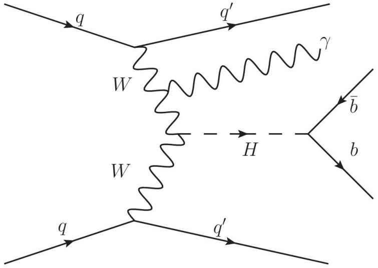


ATLAS-CONF-2020-058/

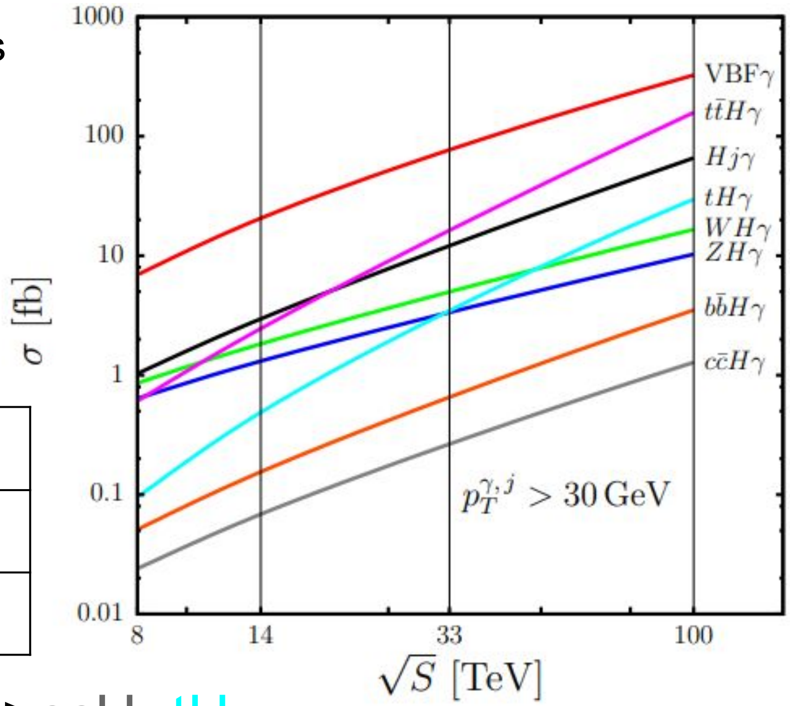
https://doi.org/10.1007/JHEP07(2016)003

VBF+ɣ

- ɣ decreases cross-section, but dramatically increases VBF yield relative to background
- ɣ from internal W: sizeable cross-section
- Enriched WW fusion compared to ZZ fusion



	Inclusive	+ɣ (LO, 30 GeV)
ggF	49 pb	3 fb
VBF	3.8 pb	19 fb



Without ɣ: ggF>>VBF>WH>ZH>ttH~bbH>>ccH~tH

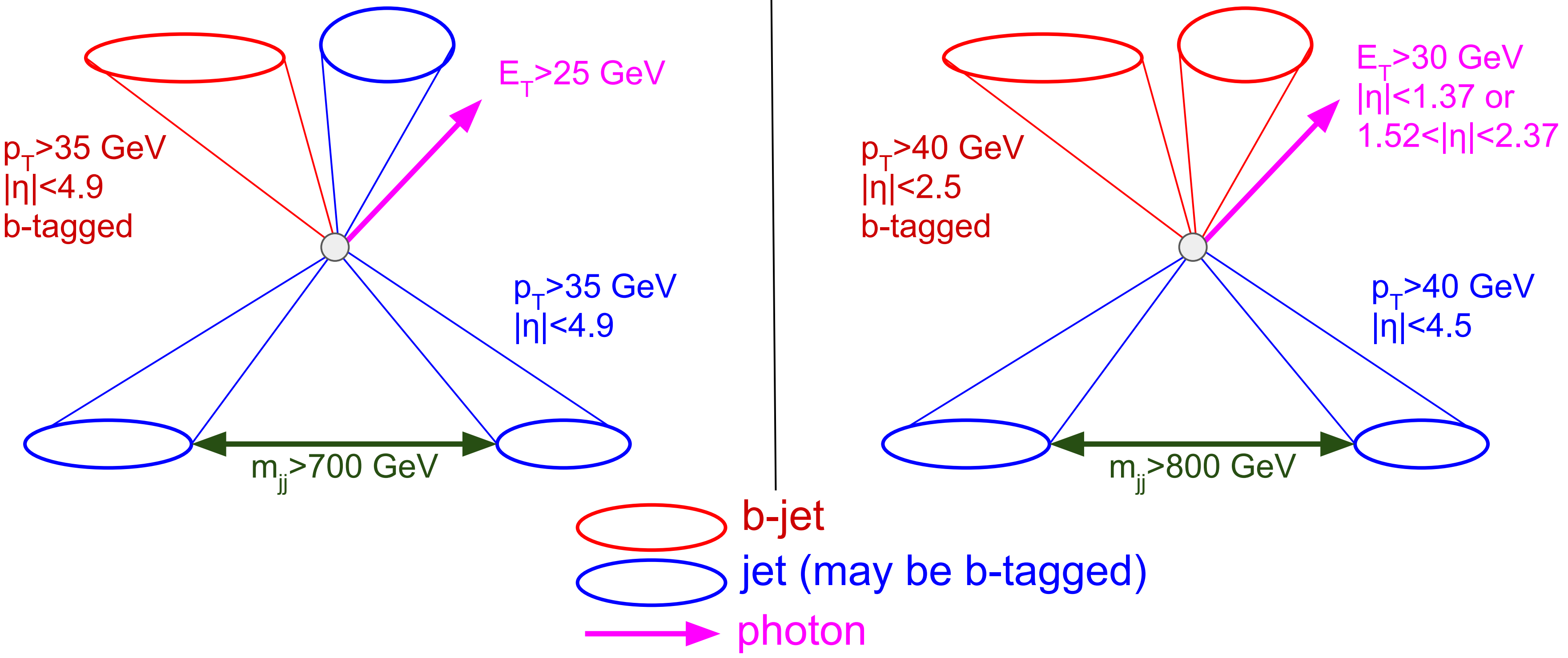
With ɣ: at 13 TeV VBF>>ggF>ttH>WH>ZH>tH>bbH>ccH

VBF+γ: Event Selection

arXiv:2010.13651

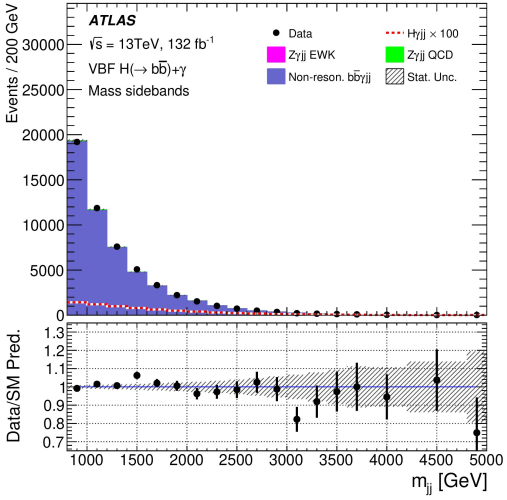
Trigger (HLT) Selection

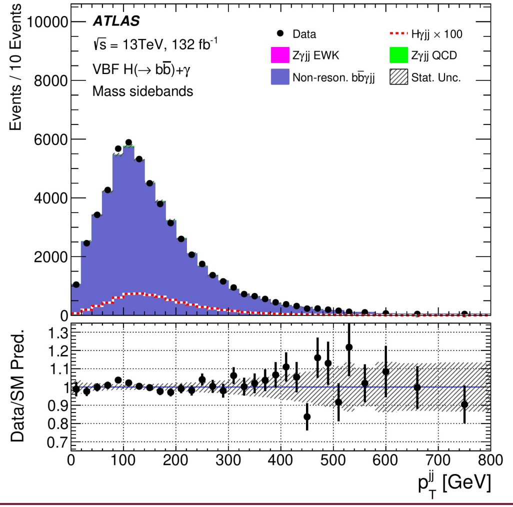
Offline Selection



Categorization

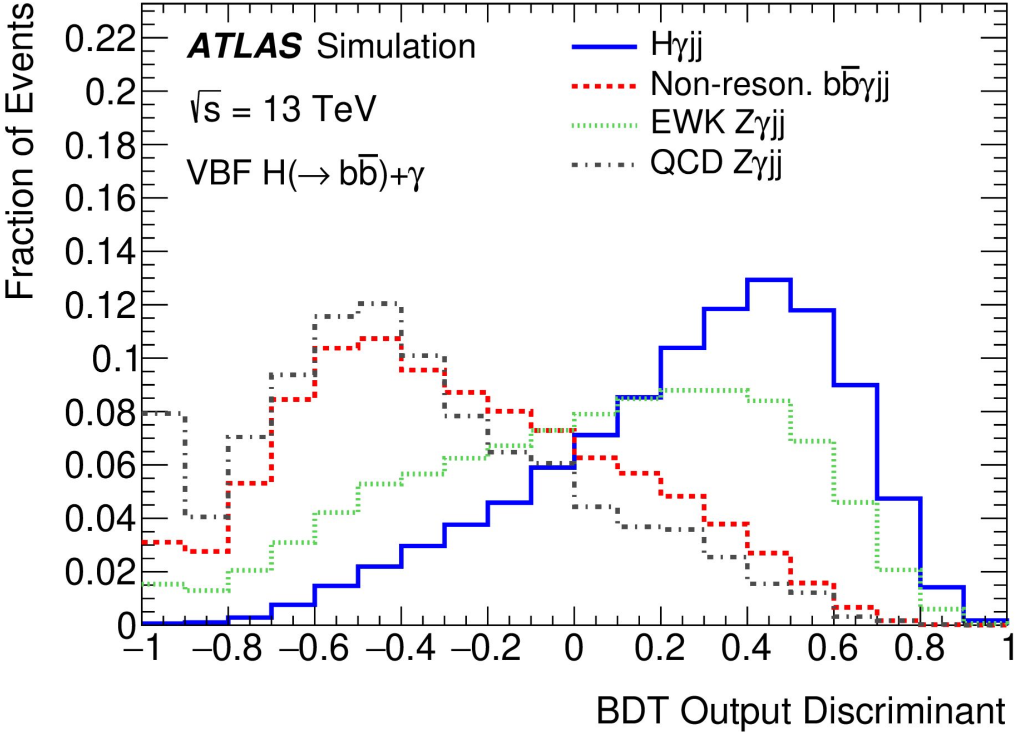
[arXiv:2010.13651](https://arxiv.org/abs/2010.13651)



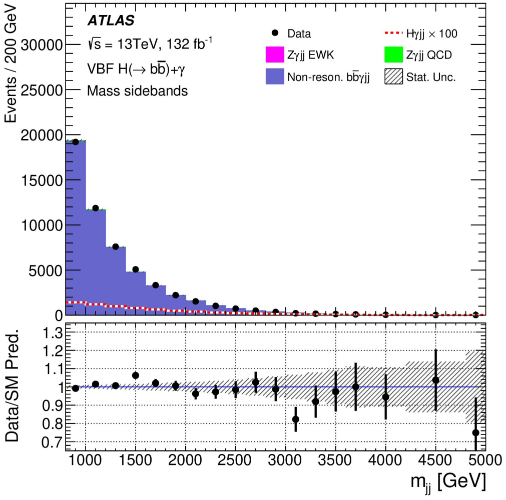


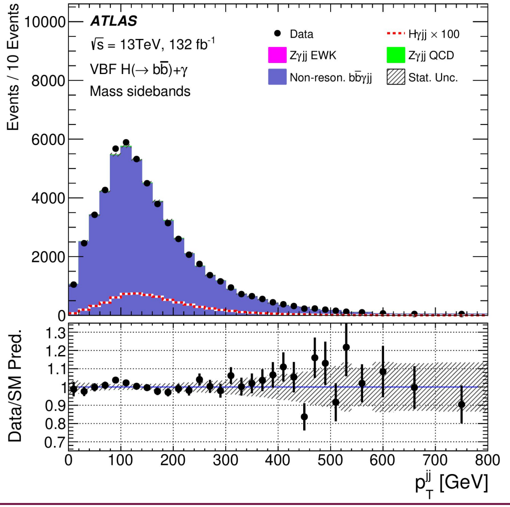
Train BDT to discriminate signal and background

Exclude variables m_{bb}-correlated, to avoid sculpting background



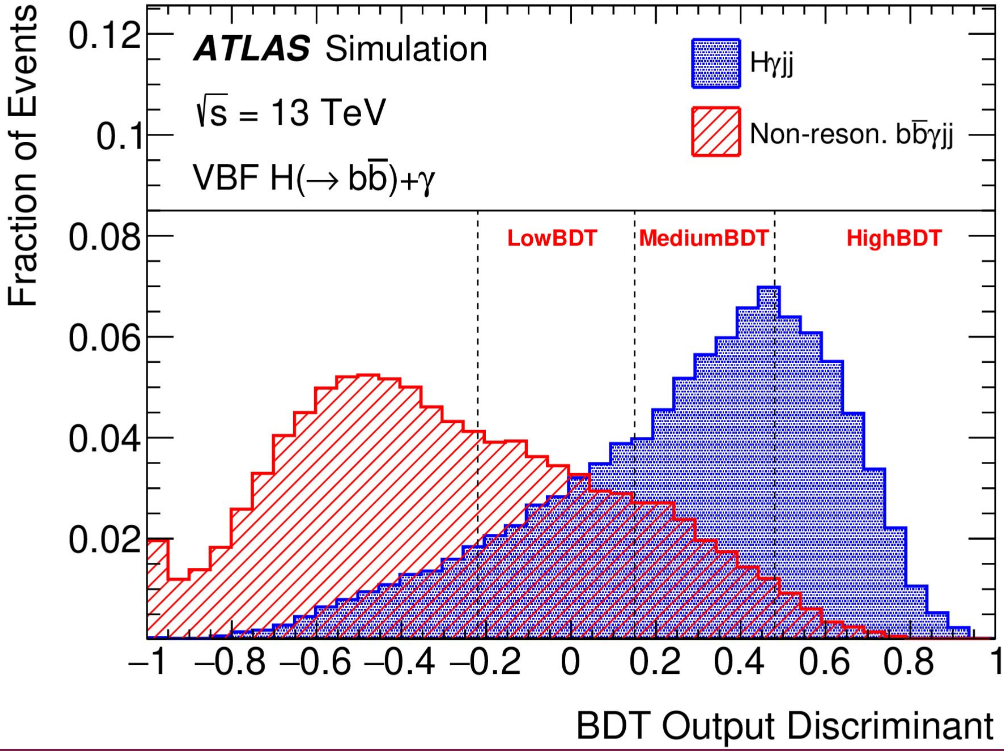
Categorization





Categorize based on score, maximizing combined significance

Fit m_{bb} distribution in each region



Model and Spurious Signal

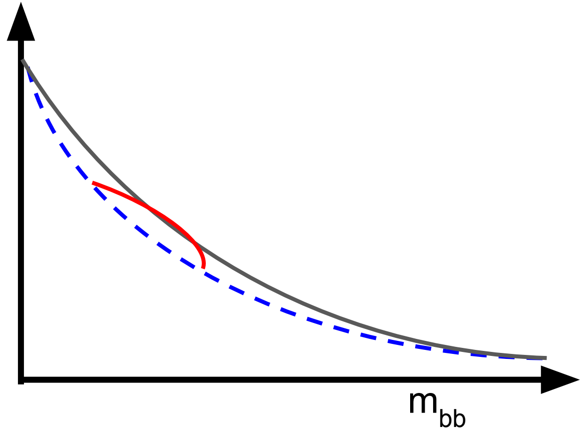
Fit data as sum of polynomial term (for non-resonant background) and peaks (Z and H)

1st order polynomial used in highest score region; 2nd order used in other regions

Primary systematic uncertainty is spurious signal - measures the potential inability of the background fit function to correctly fit the continuous background in data

Fit functions chosen to maximize sensitivity, subject to requirements on spurious signal uncertainties

Spurious signal estimated by fitting signal+background to large simulated background samples

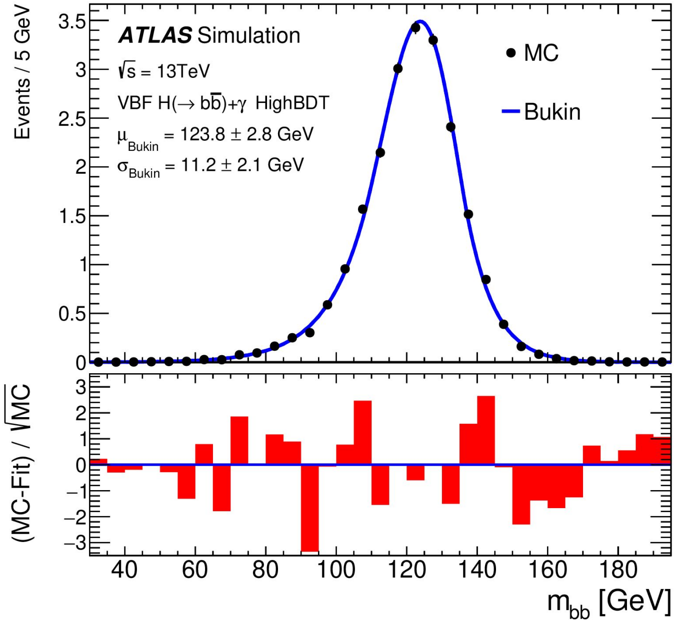


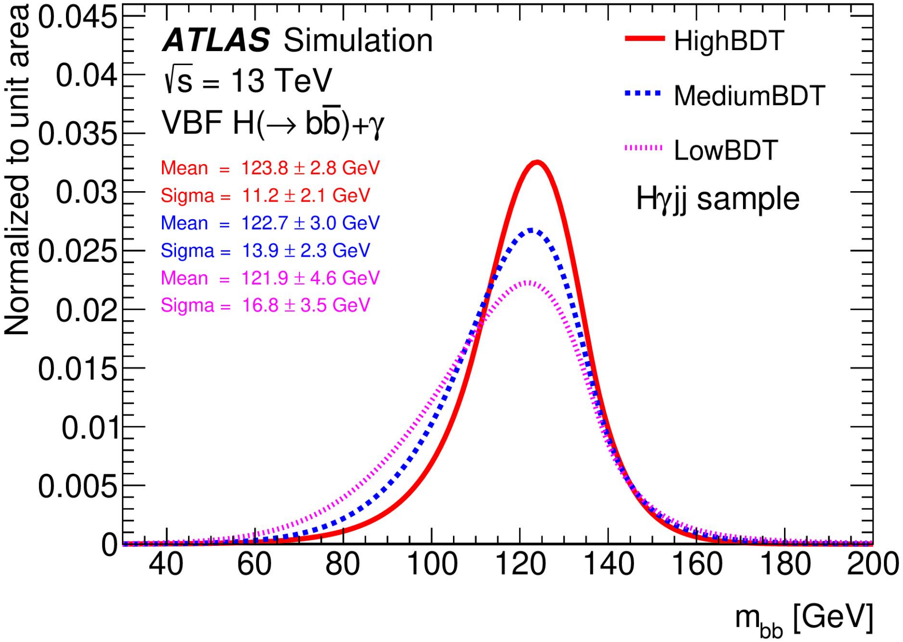
Signal and Fits

Uses b-jet energy corrections, derived in VH→bb

Signal fit of m_{bb} distribution

Parameterize Z and H with Bukin functions





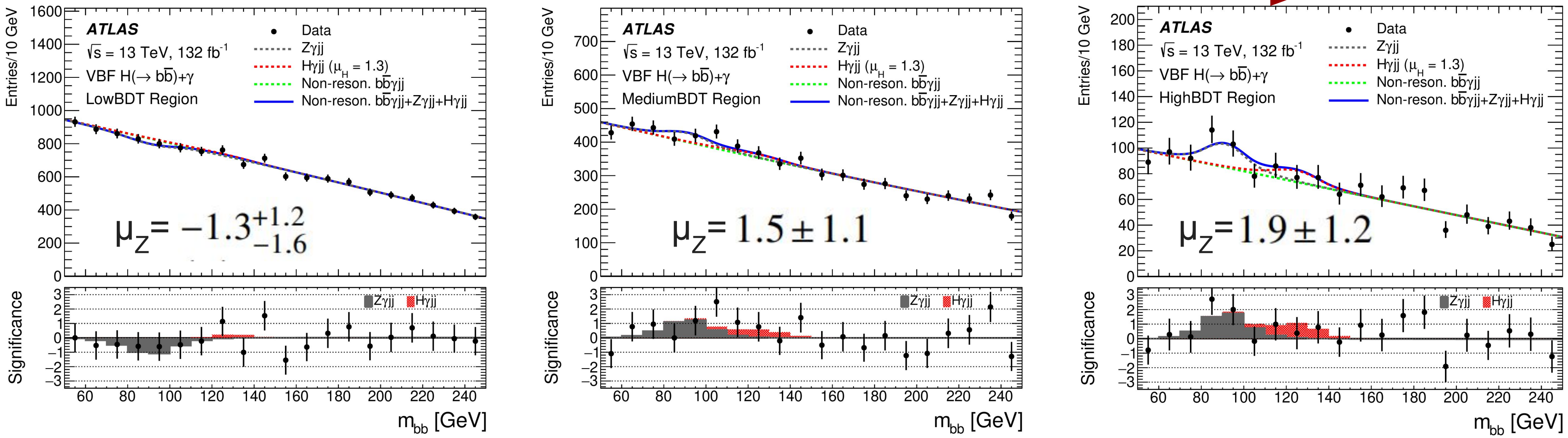
Signal+Background Fits

arXiv:2010.13651

Signal+background fits in each region

Float both Z and H components' normalizations, uncorrelating Z yields between regions

Increasing score



Uncertainties

Uncertainty	$\sigma(\mu_{H\rightarrow bb})$
Statistical	±0.31
Background shape	±0.15
Resonant background	±0.05
Theory	+0.06, -0.03
Object	+0.10, -0.05

Statistical uncertainties dominate

Uncertainties from CRs

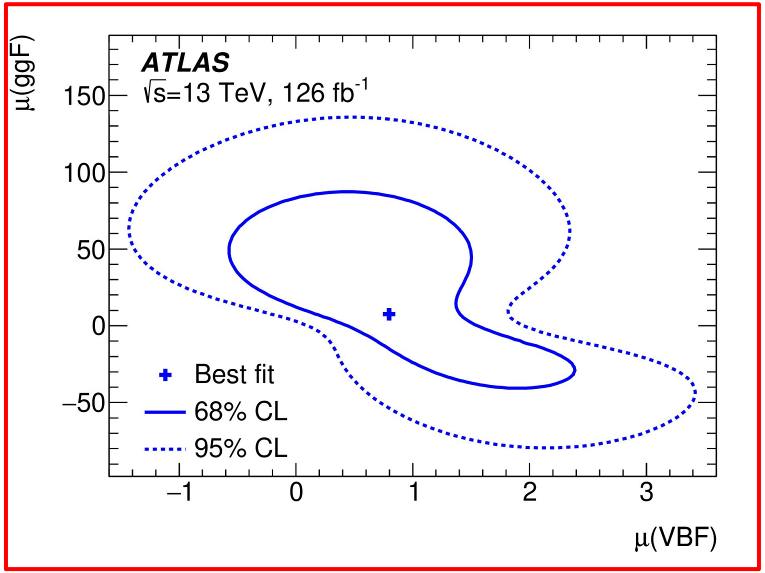
Uncertainties from data-driven $Z\rightarrow bb$ estimate

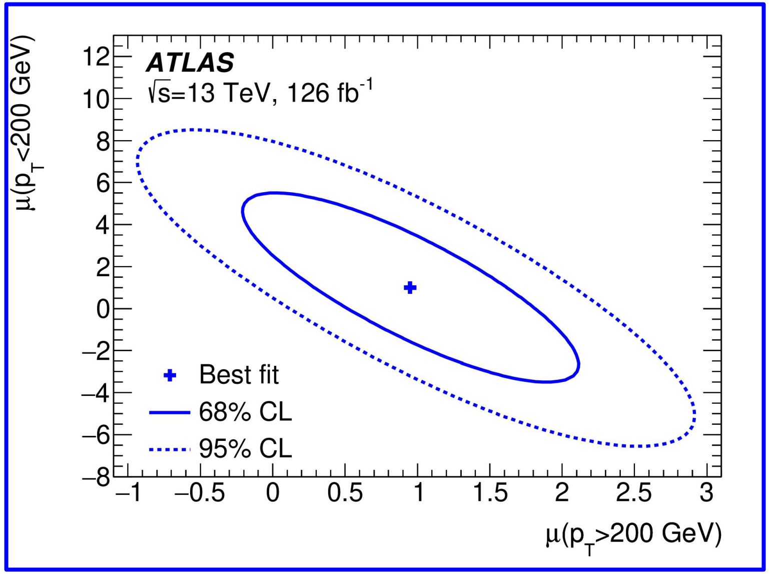
Theory uncertainties, signal only

Other uncertainties, signal only
(trigger, jet energy, b-tagging, etc.)

Results

Results	Inclusive Production	VBF Production	$p_T > 200$ GeV
Expected significance	2.9σ	2.8σ	2.3σ
Observed significance	2.7σ	2.6σ	2.2σ
Expected signal strength	$1.00^{+0.37}_{-0.36}$	$1.00^{+0.38}_{-0.37}$	$1.00^{+0.45}_{-0.43}$
Observed signal strength	$0.95^{+0.37}_{-0.35}$	$0.95^{+0.38}_{-0.36}$	$0.93^{+0.45}_{-0.43}$





arXiv:2010.13651

Results

Observe (expect)

- 1.3σ (1.0σ)
- $\mu_{H\rightarrow bb}$ = 1.3±1.0 (1.0±1.0)
- $\mu_{VBF\ H\rightarrow bb}$ = 1.3±1.0 (1.0±1.0)