$H \to b\bar{b}$ (VH, ggH, VBF)

A. de Wit on behalf of the CMS and ATLAS collaborations

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY

Outline

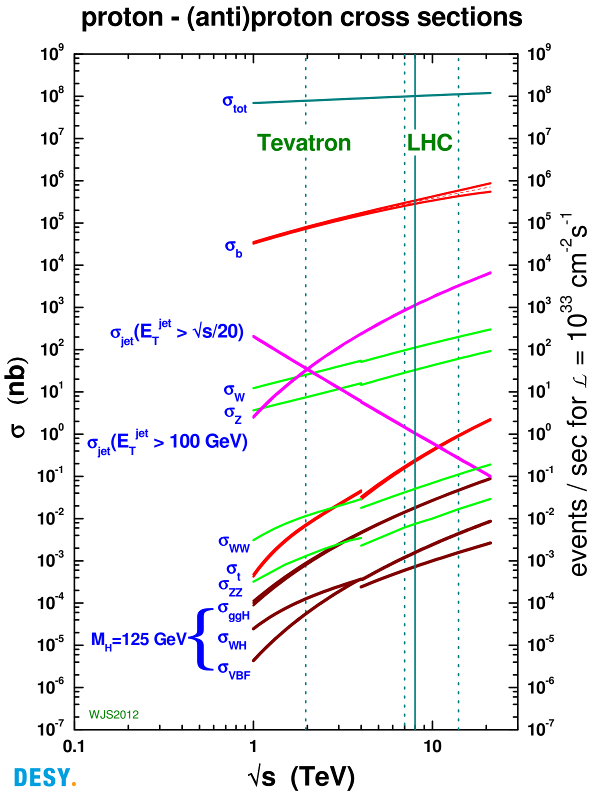
- Introduction
  - Where have we come from? What questions are we still trying to answer?

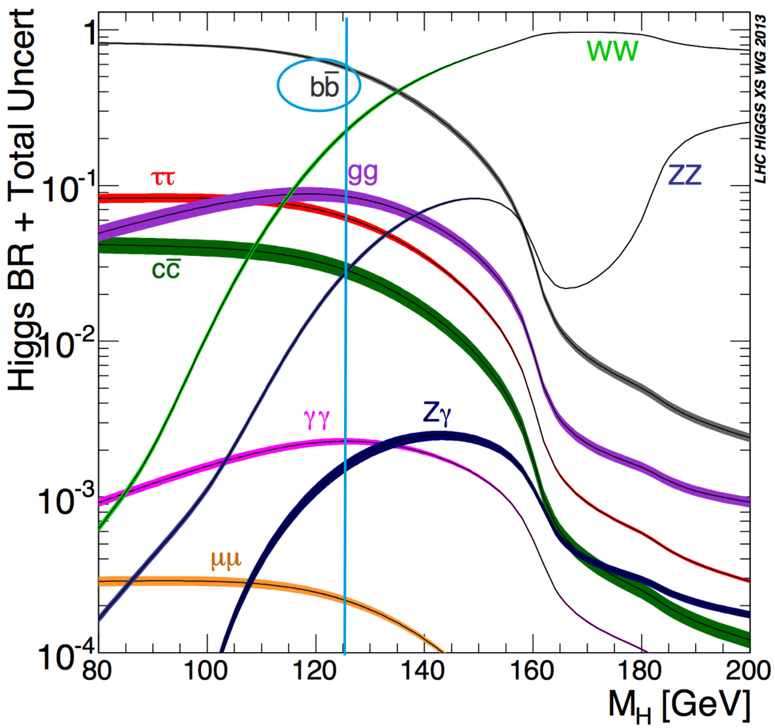
- VBF, $H \rightarrow b\bar{b}$
- ggH, $H \rightarrow b\bar{b}$
- VH, $H \rightarrow b\bar{b}$

- Summary and outlook

Introduction

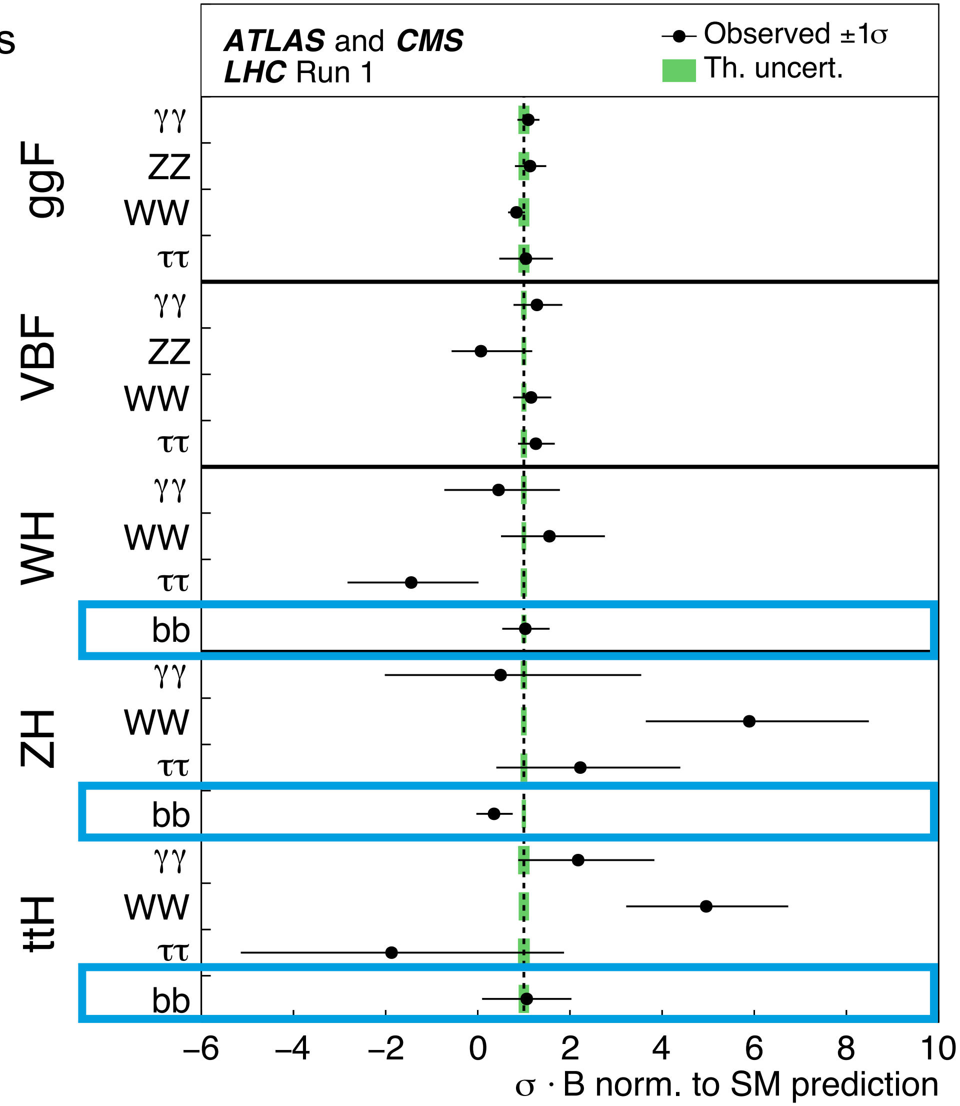
- A Higgs boson was discovered by the ATLAS and CMS collaborations in July 2012
  - The mass of the Higgs boson is known to be ~125 GeV
- H➔$b\bar{b}$: largest branching ratio: 58%
- Large multijet background makes inclusive study in this channel challenging ➔ coupling of Higgs boson to b-quarks not established yet!





Where did we come from? Where are we going?

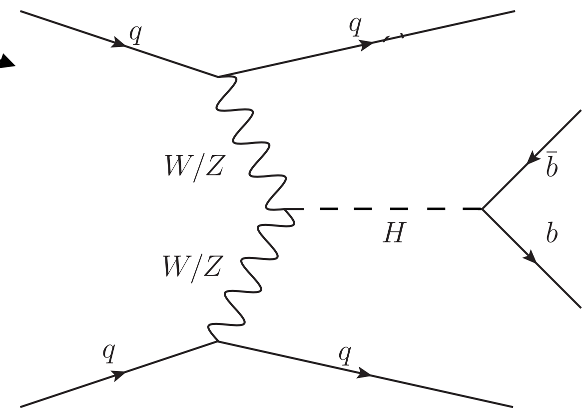
- Combining 7+8 TeV ATLAS and CMS Higgs results yielded a best-fit $\mu_{b\bar{b}}=\sigma^{*}BR/(\sigma_{SM}{}^{*}BR_{SM}) = 0.7^{+0.29}_{-0.27}$
- **Observed (expected) significance of 2.6σ (3.7σ)**
- With the growing LHC dataset, we are increasingly moving in the direction of precision Higgs measurements
  - Current focus is on firmly establishing the $H\rightarrow b\bar{b}$ decay as a stepping stone on the way
  - Precision measurements of the Higgs-b quark coupling are essential: $H\rightarrow b\bar{b}$ gives the largest contribution to the total Higgs width ➔ all other branching ratios depend on $BR(H\rightarrow b\bar{b})$
- **Precision measurements will tell us if we're dealing with the SM H boson… or if there's more!**

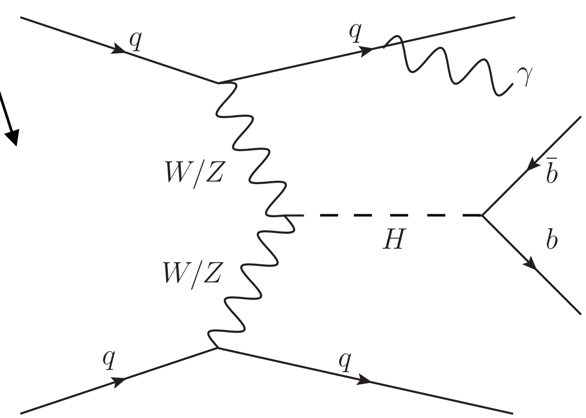


VBF $H \rightarrow b\bar{b}$

Overview

- Unique VBF topology of 2 quark jets with large rapidity gap ➔ signature for online event selection
- In ATLAS, a VBF + γ signature is employed ➔ Trade ~60 x smaller production cross section for presence of photon which can be used to cleanly trigger on and reduces backgrounds:
  - no photon radiation in gluon-gluon induced background
  - destructive interference between diagrams with photon radiation from initial- and final state quarks.
- Challenge: model the still very large multijet background

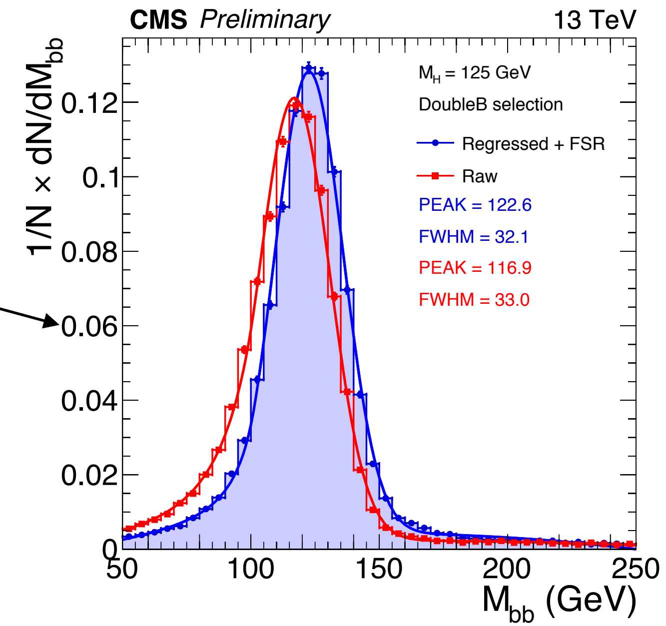


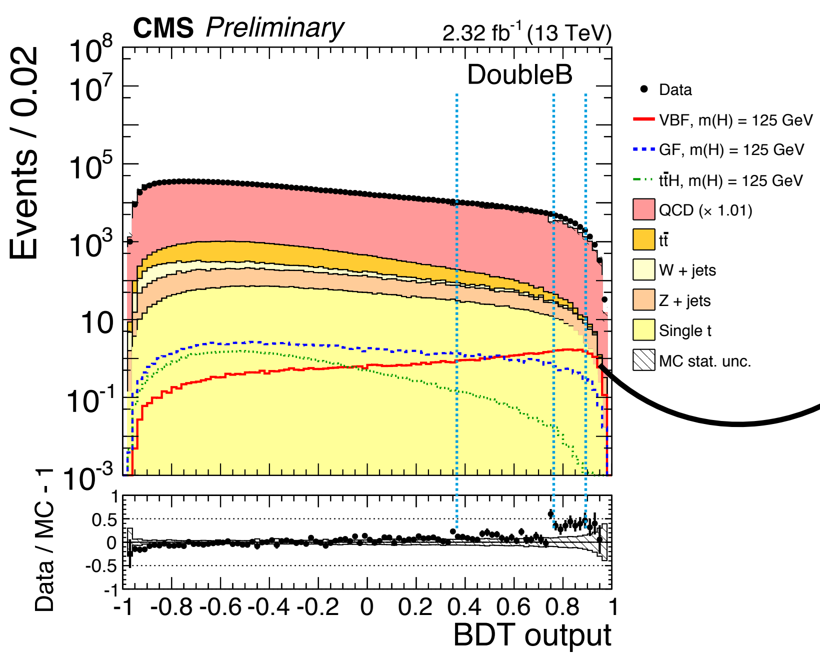


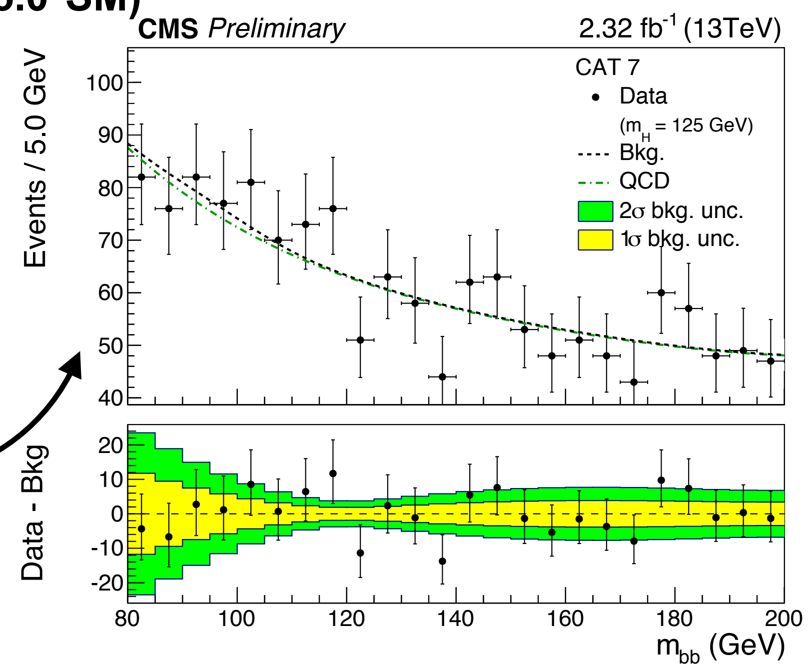
ATLAS	CMS
Partial 2016 dataset (12.6 fb^{-1})	2015 dataset (2.3 fb^{-1})

VBF - CMS

- **Target inclusive VBF topology**: selection of events with VBF topology plus b-tagged jet(s) ➔ **multi-jet trigger** challenging to maintain in high-PU environment
- **Improve** m_{bb} resolution using multivariate regression techniques
- **Categorisation**: use BDTs trained with variables that are weakly correlated with bb kinematics and use the BDT output to define categories
- **Fit** the m_{bb} distribution simultaneously in all categories to extract signal (major backgrounds fit to data analytically)
- **Result: observed (expected) upper limit 3.0*SM (5.0*SM)**

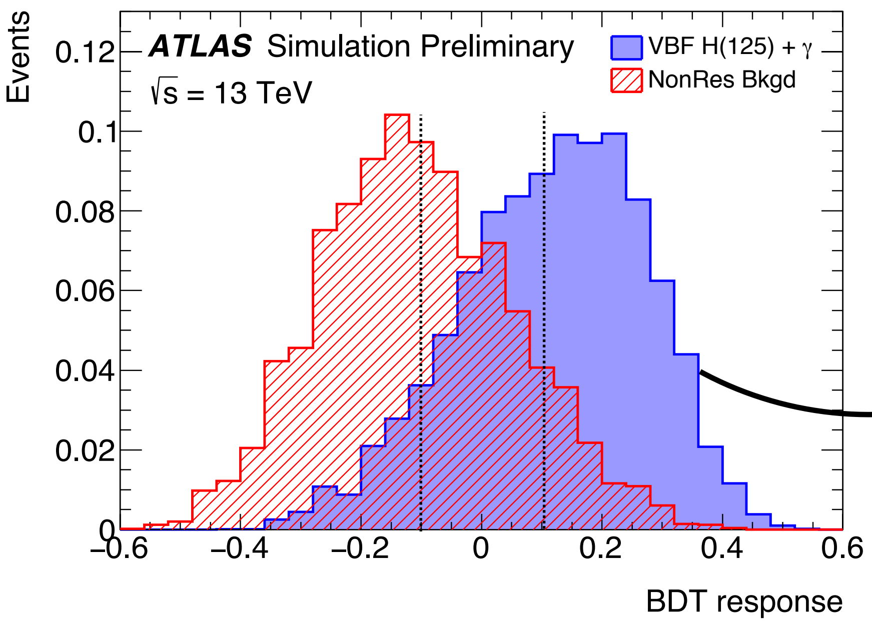


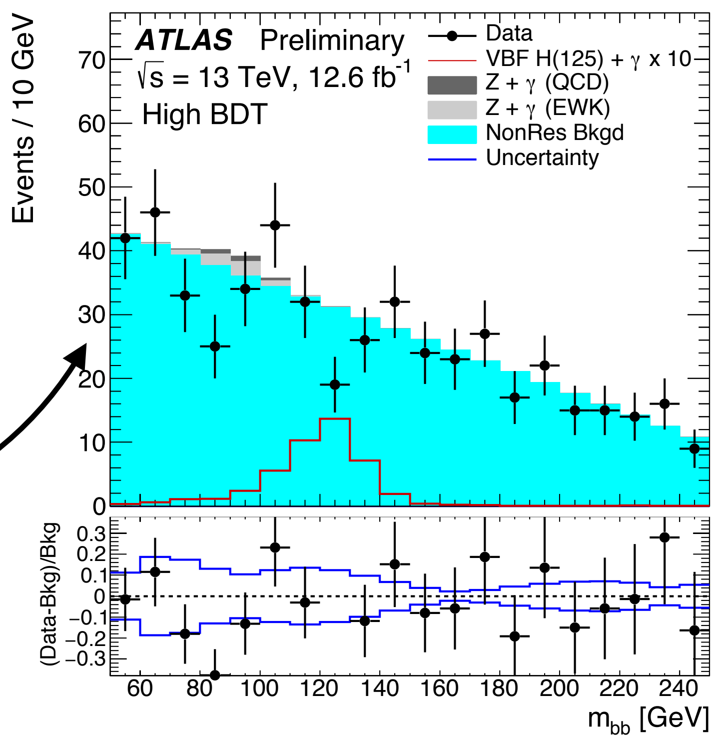




VBF - ATLAS

- **Target VBF+γ topology**: online selection of a photon and at least 4 jets. Offline require 2 central b-tagged jets and 2 jets consistent with VBF signature
- **Categorisation:** use a BDT to separate signal and background, ensuring that the BDT variables are weakly correlated with the kinematics of the bb system
- **Fit** the m_{bb} distribution in all categories simultaneously (major backgrounds fit to data analytically)
- **Result: observed (expected) upper limit 4.0*SM (6.0*SM)**
- Setup also used to search for $Z(b\bar{b})$+γjj production: observed (expected) upper limit 2.0*SM (1.8*SM)



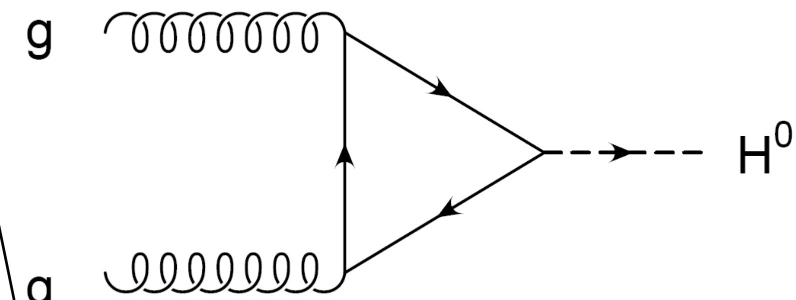


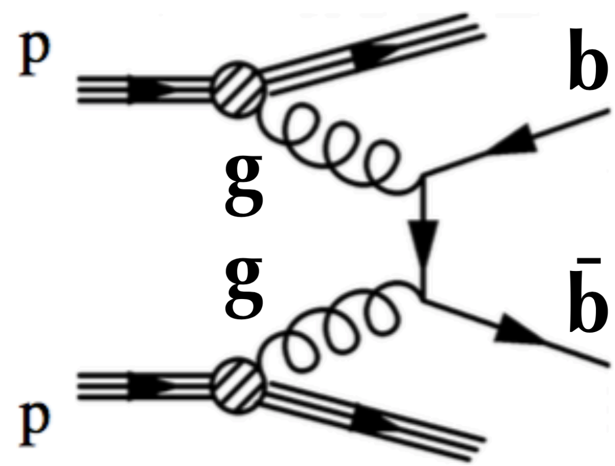
ggH, H→b$\bar{b}$

Overview

- Largest production cross section, but suffers from overwhelming background from heavy flavour multijet production
  - Up to recently a search for ggH production in the H→b$\bar{b}$ decay channel would have been deemed impossible…
- At high H p_T the two b-jets are likely to merge into a single 'fat' jet ➔ **exploit di-b jet substructure to make an inclusive H→b$\bar{b}$ search at high H p_T possible**

ATLAS	CMS
-	**Full 2016 dataset ($35.9\ fb^{-1}$)**



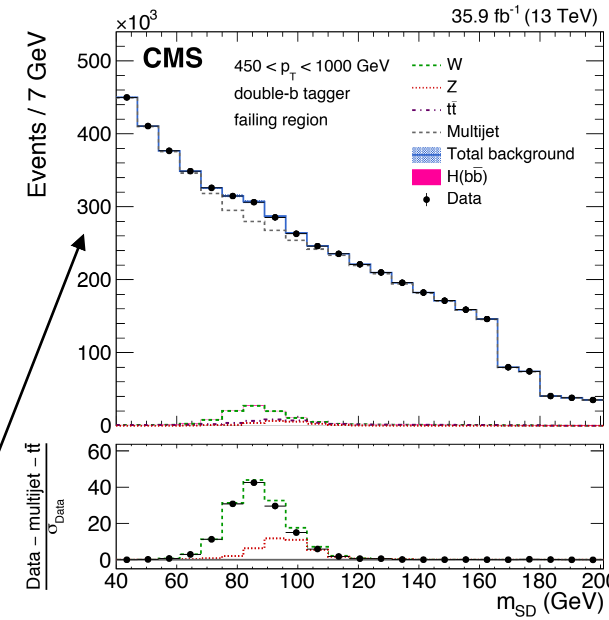


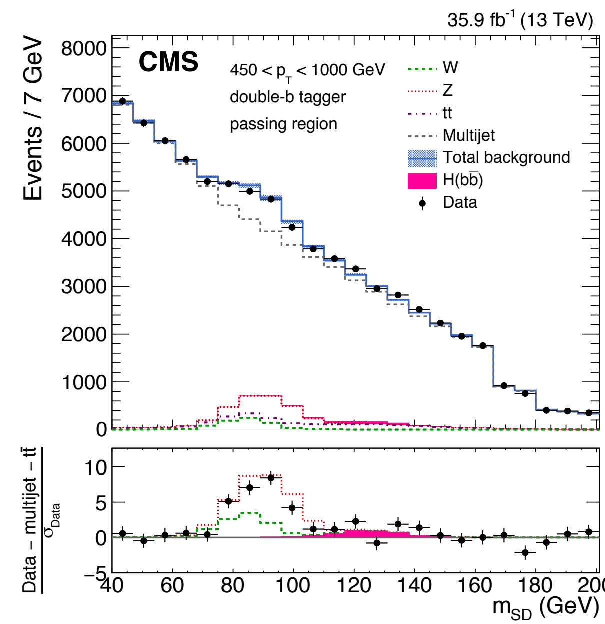
PRL 120(2018) 071802

Inclusive (ggH) H→b$\bar{b}$

- Require a high p_T (>450 GeV) wide-cone jet
  - Exploit **2-prong jet substructure** and **b-tagging information of the sub-jets** to reduce multijet background

    ➔ largest acceptance (75%) for ggH production, but other modes taken into account too
  - Remove soft radiation from candidate jet to provide better separation between signal and multijet background with the jet mass shape
- Major backgrounds from multijet production, with smaller contributions from tt, W, and Z production
  - Multijet background estimated from a data sample with inverted double-b tag requirement in a simultaneous fit with the signal region
- Validation of the method: extract Z(b$\bar{b}$) production cross section times branching ratio:
  - $\mu_Z = 0.78 \pm 0.14\,(\text{stat})\,^{+0.19}_{-0.13}\,(\text{syst})$ ➔ observed (expected) significance of 5.1σ (5.8σ). **First observation of Z(b$\bar{b}$) in single-jet topology!**
- **Results**:
  $\mu_H = 2.3 \pm 1.5\,(\text{stat.})\,^{+1.0}_{-0.4}\,(\text{syst.})$ ➔ observed (expected) significance of 1.5σ(0.7σ)

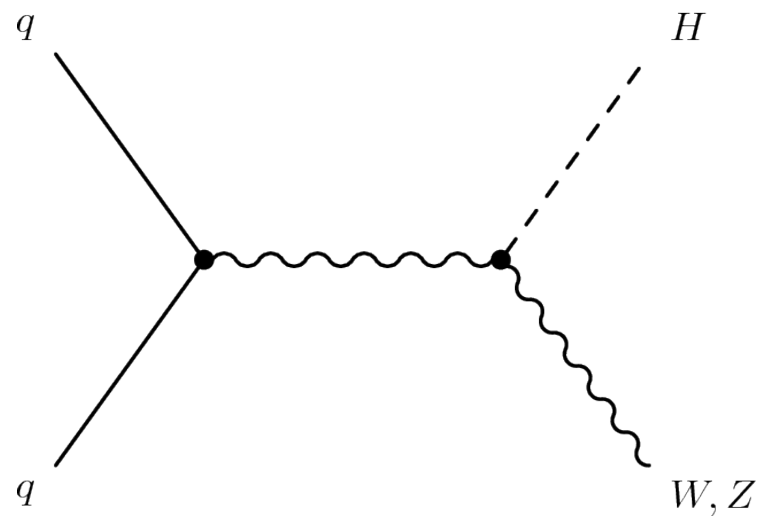


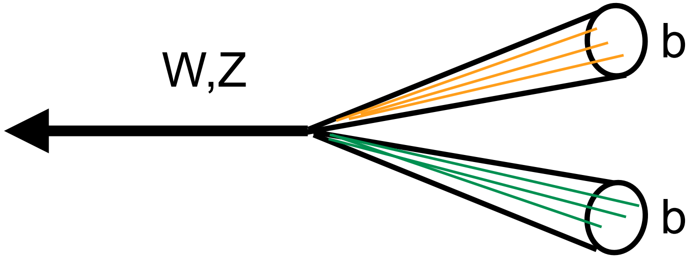


VH,H→b$\bar{b}$

Overview

- Higgs boson produced in association with a vector boson
  - Leptonically decaying vector boson gives a clean signature to tag ➔ helpful for online selection
  - Reduced background from multijet production
- Most sensitive channel for H➔b$\bar{b}$ studies despite smaller cross-section than gluon fusion and VBF production





ATLAS	CMS
Full 2015+2016 dataset (36.1 fb^{-1})	**Full 2016 dataset (35.9 fb^{-1})**

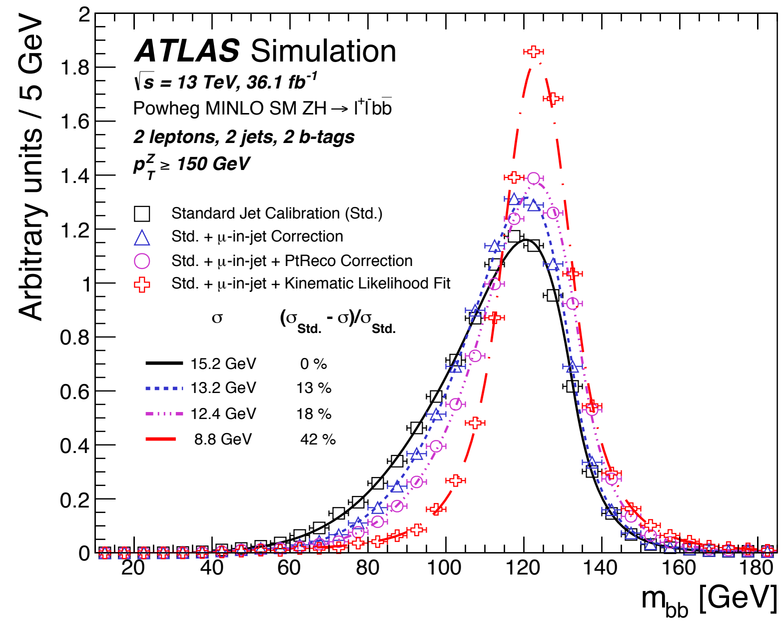
JHEP 12 (2017) 024

VH - ATLAS

Analysis strategy

- **Select** events with 0, 1 or 2-leptons (e/μ), consistent with W/Z decay, and 2 b-tagged jets
- **Categorisation** based on V p_T and number of jets
- **Improve** the m_{bb} resolution using jet corrections
  - **μ-in-jet:** accounts for b/c→μ decays not depositing full energy in calorimeter
  - **PtReco:** correction to jet response based on difference between reconstructed b-jets and MC truth jets
  - **Kinematic fit (2 lepton channel):** exploit transverse momentum balance of the ZH(llb$\bar{b}$) decay to improve jet resolution
- Use BDTs to increase the separation between signal and background
  - Variables include V p_T, m_{bb}, jet kinematics,…
- **Fit** for signal using the BDT output as final discriminant, simultaneously fitting control regions to constrain some of the backgrounds
- **Validate** the analysis strategy using a di-boson analysis and an analysis using m_{bb} as discriminating variable

	0-lepton		1-lepton		2-lepton			
V p_T	> 150 GeV		> 150 GeV		75-150 GeV		> 150 GeV	
Njets	2	3	2	3	2	≥3	2	≥3



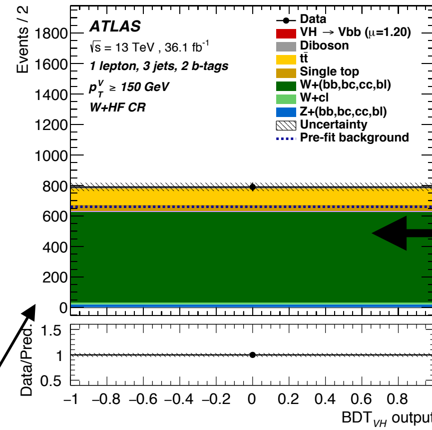
JHEP 12 (2017) 024

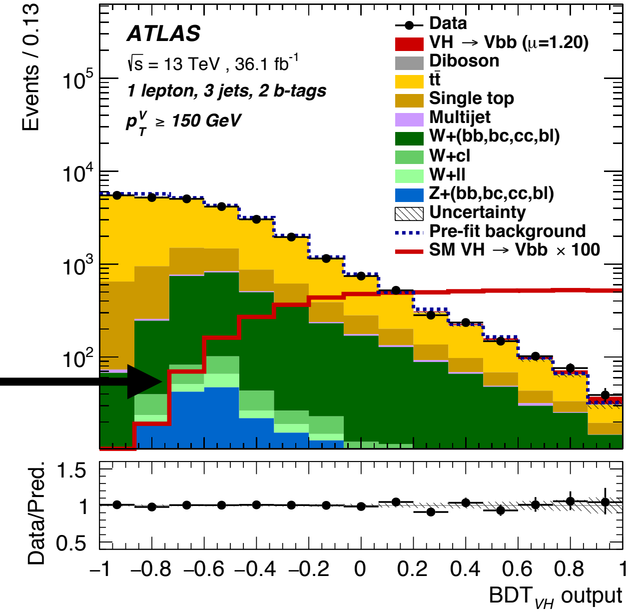
VH - ATLAS

Background normalisation

- **Background normalisation** constrained in the fit for major backgrounds ($t\bar{t}$, Z+heavy flavour, W+heavy flavour)

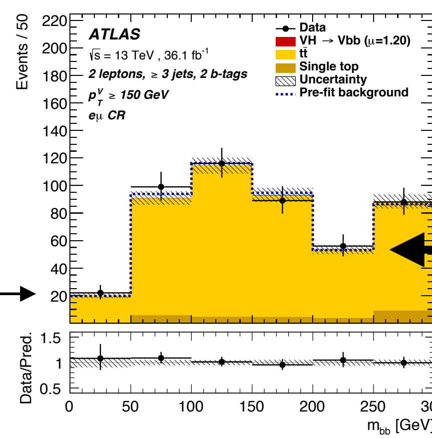
- **In 1-lepton channel**, constrain the **W+HF normalisation** via control regions

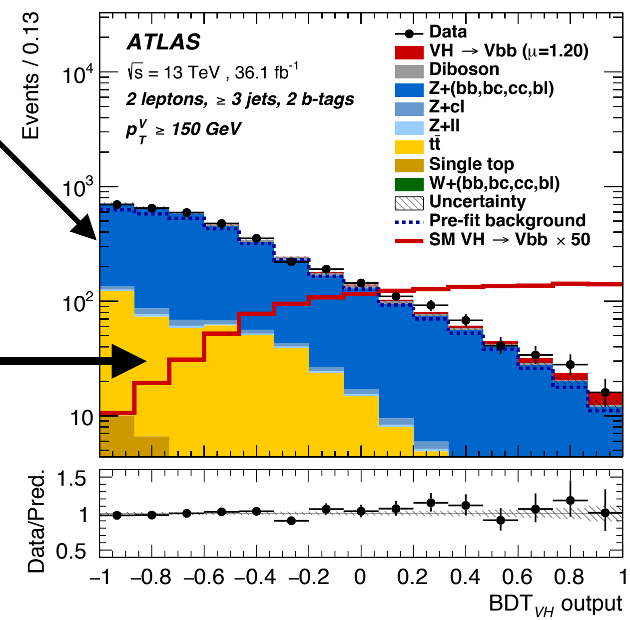




- **Z+HF normalisation** constrained by the **2-lepton channel signal region**

- **In 2-lepton channel**, constrain the **$t\bar{t}$ normalisation** via control regions





VH - ATLAS

Results

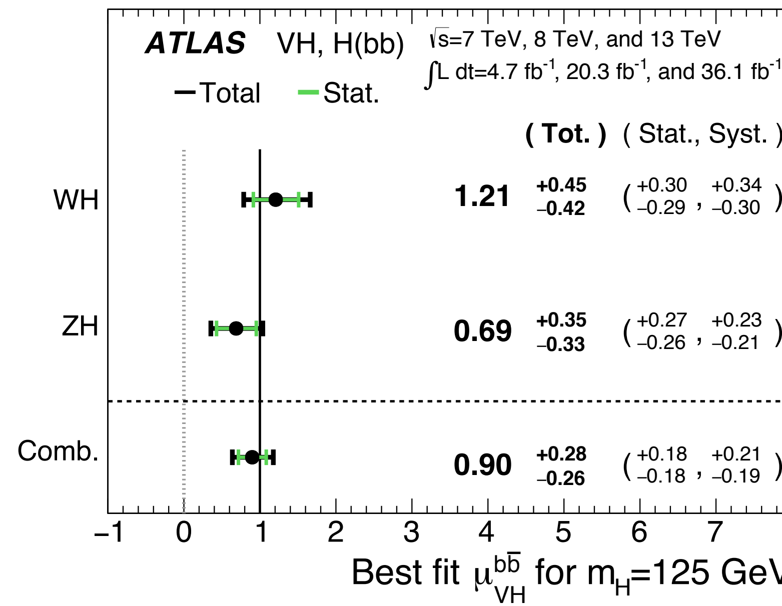
- Best-fit signal strength $\mu = 1.20^{+0.24}_{-0.23}$ (stat) $^{+0.34}_{-0.28}$(syst) ➔ observed (expected) significance 3.5σ (3.0σ)
- Combining the results with Run 1: $\mu=0.9\pm0.18$(stat.) $^{+0.21}_{-0.19}$(syst.) ➔ **observed (expected) significance 3.6σ (4.0σ)**

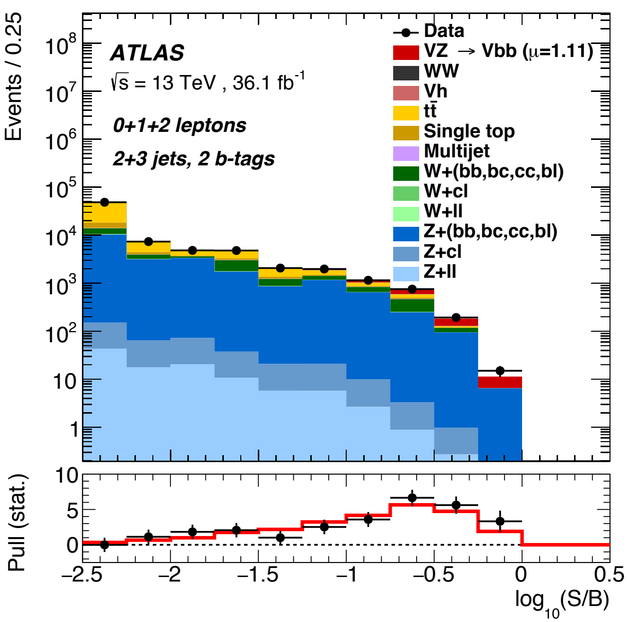
➔Results compatible with SM H expectation

- Cross check results:

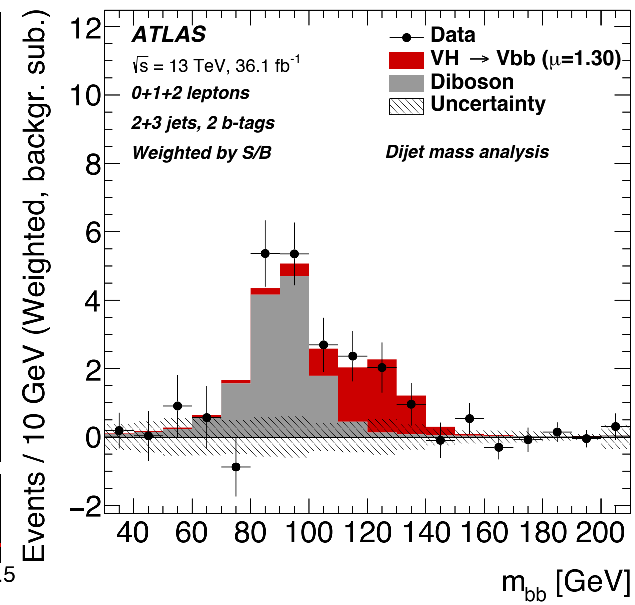
VZ($b\bar{b}$) analysis: $\mu_{VZ}=1.11^{+0.12}_{-0.11}$ (stat.) $^{+0.22}_{-0.19}$ (syst.)

m_{bb} analysis: $\mu=1.30^{+0.28}_{-0.27}$(stat.) $^{+0.37}_{-0.29}$ (syst.)

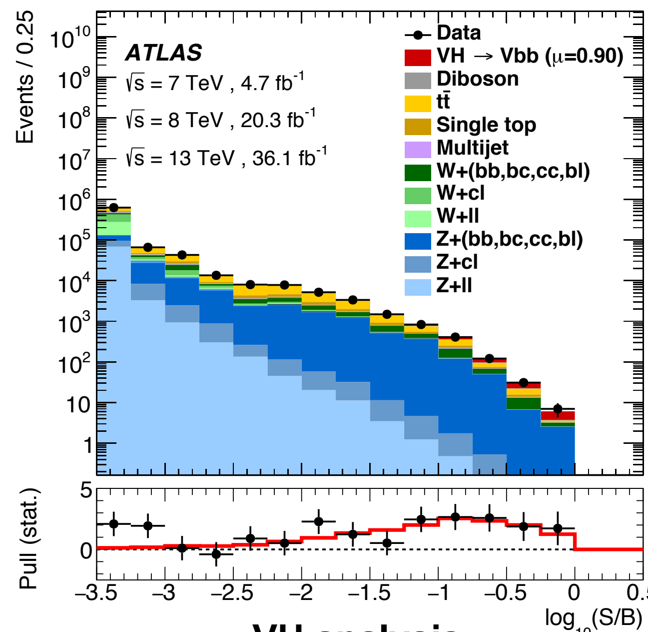




VZ analysis



m_{bb} analysis



VH analysis

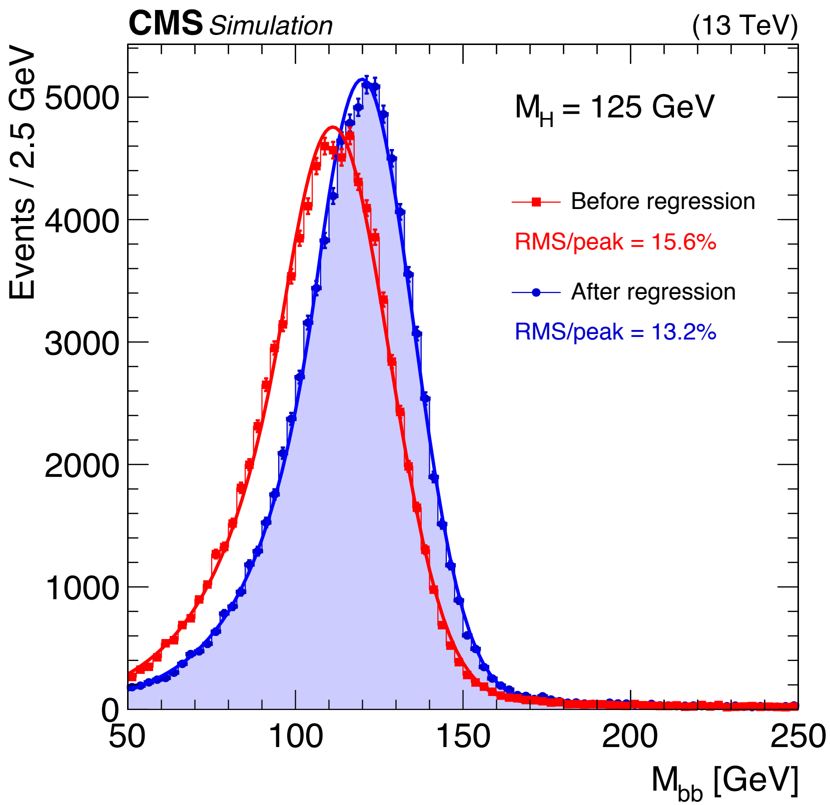
PLB 780 (2018) 501-532

VH - CMS

Strategy

- **Select** events with 0, 1 or 2-leptons (e/μ), consistent with W/Z decay, and 2 b-tagged jets
- **Categorisation** based on V-boson p_T
- **Improve** the m_{bb} resolution using multivariate regression techniques
- Use BDTs to increase the separation between signal and background
  - Variables include V p_T, m_{bb}, other jet and di-jet kinematics,…
- **Fit** for signal using the BDT output as final discriminant, simultaneously fitting control regions to constrain some of the backgrounds
- **Validate** the analysis strategy using a di-boson analysis

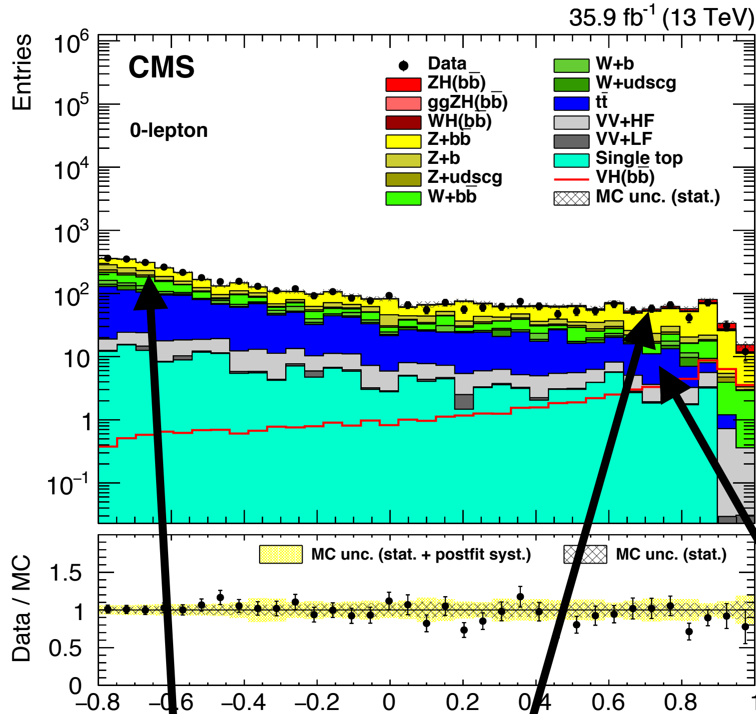
	0-lepton	1-lepton	2-lepton	
V p_T	> 170 GeV	> 100 GeV	50-150 GeV	> 150 GeV



VH - CMS

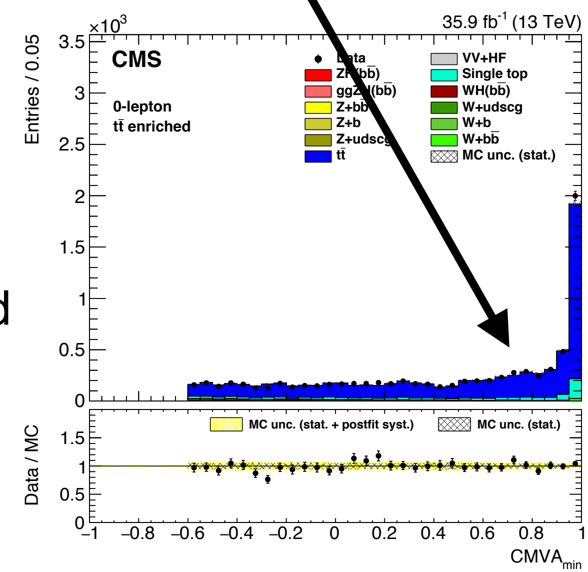
Background normalisation

- **Background normalisation** constrained in the fit for major backgrounds ($t\bar{t}$, Z+heavy flavour, Z+light flavour, W+heavy flavour, W+light flavour)
- **Always constrained via control regions**
- In control regions, fit the b-tagging discriminant of the least b-tagged signal jet



Z+light flavour in dedicated Z+LF CR

Z+bb in dedicated Z+HF CR

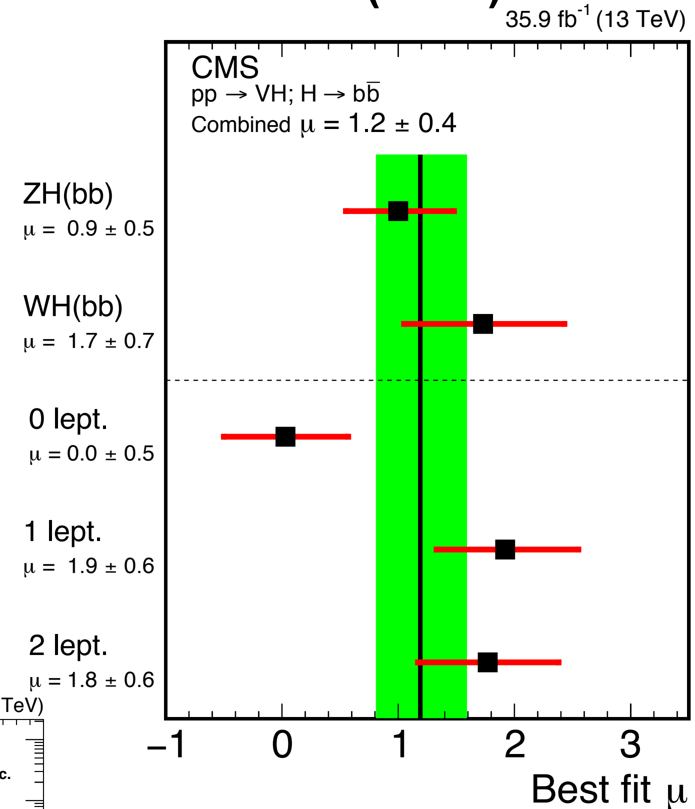


PLB 780 (2018) 501-532

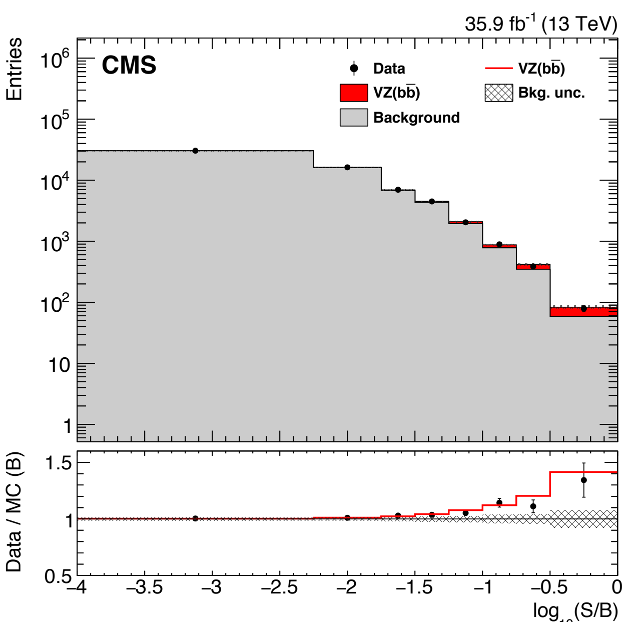
VH - CMS

Results

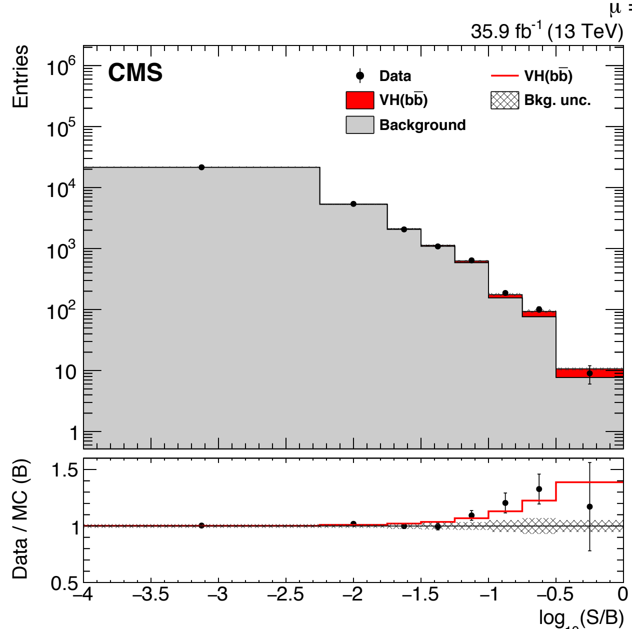
- Best-fit signal strength $\mu = 1.19^{+0.21}_{-0.20}\,(\text{stat})\,^{+0.34}_{-0.32}\,(\text{syst})$
  ➔ observed (expected) significance 3.3σ (2.8σ)
- Combining the results with Run 1:
  $\mu = 1.06^{+0.31}_{-0.29}$ (stat+syst.) ➔ **observed (expected) significance 3.8σ (3.8σ)**
- VZ cross-check analysis:
  $\mu_{VZ} = 1.02 \pm 0.22$ (stat+syst)



Results compatible with SM H expectation



VZ analysis



VH analysis

JHEP 12 (2017) 024
PLB 780 (2018) 501-532

VH - uncertainty contributions

ATLAS

Source of uncertainty		σ_μ
Total		0.39
Statistical		0.24
Systematic		0.31
Experimental uncertainties		
Jets		0.03
E_T^{miss}		0.03
Leptons		0.01
b-tagging	b-jets	0.09
	c-jets	0.04
	light jets	0.04
	extrapolation	0.01
Pile-up		0.01
Luminosity		0.04
Theoretical and modelling uncertainties		
Signal		0.17
Floating normalisations		0.07
Z + jets		0.07
W + jets		0.07
$t\bar{t}$		0.07
Single top quark		0.08
Diboson		0.02
Multijet		0.02
MC statistical		0.13

CMS

Source	Type	Individual contribution to the μ uncertainty (%)	Effect of removal to the μ uncertainty (%)
Scale factors ($t\bar{t}$, V+jets)	norm.	9.4	3.5
Size of simulated samples	shape	8.1	3.1
Simulated samples' modeling	shape	4.1	2.9
b tagging efficiency	shape	7.9	1.8
Jet energy scale	shape	4.2	1.8
Signal cross sections	norm.	5.3	1.1
Cross section uncertainties (single-top, VV)	norm.	4.7	1.1
Jet energy resolution	shape	5.6	0.9
b tagging mistag rate	shape	4.6	0.9
Integrated luminosity	norm.	2.2	0.9
Unclustered energy	shape	1.3	0.2
Lepton efficiency and trigger	norm.	1.9	0.1

- In both ATLAS and CMS a large component of the uncertainty on the measured signal strength is statistical
- Large systematic uncertainties
  - Main contributions to systematic uncertainty not always the same between two experiments

Summary & outlook

- $H \rightarrow b\bar{b}$ decays are actively being studied in the VBF, gluon fusion and VH production modes at the LHC
- Results discussed:

	95% upper limit obs(exp)	
	ATLAS	**CMS**
VBF (13 TeV)	<4.0*SM (<6.0*SM)	<3.0*SM (<5.0*SM)

	Best-fit μ		Significance obs(exp)	
	ATLAS	**CMS**	**ATLAS**	**CMS**
ggH (13 TeV)	-	2.3 ± 1.5 (stat) $^{+1.0}_{-0.4}$ (syst.)	-	1.5σ (0.7σ)
VH (13 TeV)	$1.20^{+0.24}_{-0.23}$ (stat) $^{+0.34}_{-0.28}$ (syst)	$1.19^{+0.21}_{-0.20}$ (stat) $^{+0.34}_{-0.32}$ (syst)	3.5σ (3.0σ)	3.3σ (2.8σ)
VH (Run1&2)	0.90 ± 0.18 (stat) $^{+0.21}_{-0.19}$ (syst)	$1.06^{+0.31}_{-0.29}$ (stat+syst)	3.6σ (4.0σ)	3.8σ (3.8σ)

- Results from ATLAS and CMS compatible with each other, best-fit signal strengths compatible with the SM, μ=1
- An additional ~40 fb^{-1} already collected during the 2017 LHC run, and the machine is just starting up ready for the 2018 run
  - Exciting times ahead for studies in the $H \rightarrow b\bar{b}$ decay mode, with precision measurements visible on the horizon

Backup

VH

BDT inputs

ATLAS

Variable	0-lepton	1-lepton	2-lepton
p_T^V	$\equiv E_T^{miss}$	$\times$	$\times$
E_T^{miss}	$\times$	$\times$	$\times$
$p_T^{b_1}$	$\times$	$\times$	$\times$
$p_T^{b_2}$	$\times$	$\times$	$\times$
m_{bb}	$\times$	$\times$	$\times$
$\Delta R(\vec{b}_1, \vec{b}_2)$	$\times$	$\times$	$\times$
$\lvert\Delta\eta(\vec{b}_1, \vec{b}_2)\rvert$	$\times$		
$\Delta\phi(\vec{V}, \vec{bb})$	$\times$	$\times$	$\times$
$\lvert\Delta\eta(\vec{V}, \vec{bb})\rvert$			$\times$
m_{eff}	$\times$		
$\min[\Delta\phi(\vec{\ell}, \vec{b})]$		$\times$	
m_T^W		$\times$	
$m_{\ell\ell}$			$\times$
m_{top}		$\times$	
$\lvert\Delta Y(\vec{V}, \vec{bb})\rvert$		$\times$	
	Only in 3-jet events		
$p_T^{jet_3}$	$\times$	$\times$	$\times$
m_{bbj}	$\times$	$\times$	$\times$

CMS

Variable	Channels utilizing
$M(jj)$: dijet invariant mass	All
$p_T(jj)$: dijet transverse momentum	All
$p_T(V)$: vector boson transverse momentum	All
$CMVA_{max}$: value of CMVA for the Higgs boson daughter with largest CSV value	2-lepton, 0-lepton
$CMVA_{min}$: value of CMVA for the Higgs boson daughter with second largest CSV value	All
$CMVA_{add}$: value of CMVA for the additional jet with largest CSV value	0-lepton
$\Delta\phi(V, H)$: azimuthal angle between V and dijet	All
$p_T(j)$: transverse momentum of each Higgs boson daughter	2-lepton, 0-lepton
$p_T(add.)$: transverse momentum of leading additional jet	0-lepton
$\lvert\Delta\eta(jj)\rvert$: difference in η between Higgs boson daughters	2-lepton, 0-lepton
$\Delta R(jj)$: distance in η–ϕ between Higgs boson daughters	2-lepton
N_{aj}: number of additional jets N.B. definition slightly different per channel	1-lepton, 2-lepton
$p_T(jj)/p_T(V)$: p_T balance between Higgs boson candidate and vector boson	2-lepton
: Z boson mass	2-lepton
SA5: number of soft activity jets with $p_T > 5$ GeV	All
M_t: reconstructed top quark mass	1-lepton
$\Delta\phi(E_T^{miss}, \ell)$: azimuthal angle between E_T^{miss} and lepton	1-lepton
E_T^{miss}: missing transverse energy	1-lepton, 2-lepton
$m_T(W)$: W transverse mass	1-lepton
: difference in ϕ between Higgs boson daughters	0-lepton
$\Delta\phi(E_T^{miss}, jet.)$: azimuthal angle between E_T^{miss} and the closest jet with $p_T > 30$ GeV	0-lepton