$H\rightarrow b\bar{b}$ and $H\rightarrow c\bar{c}$ measurements from CMS

A. de Wit (DESY) on behalf of the CMS collaboration

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Outline

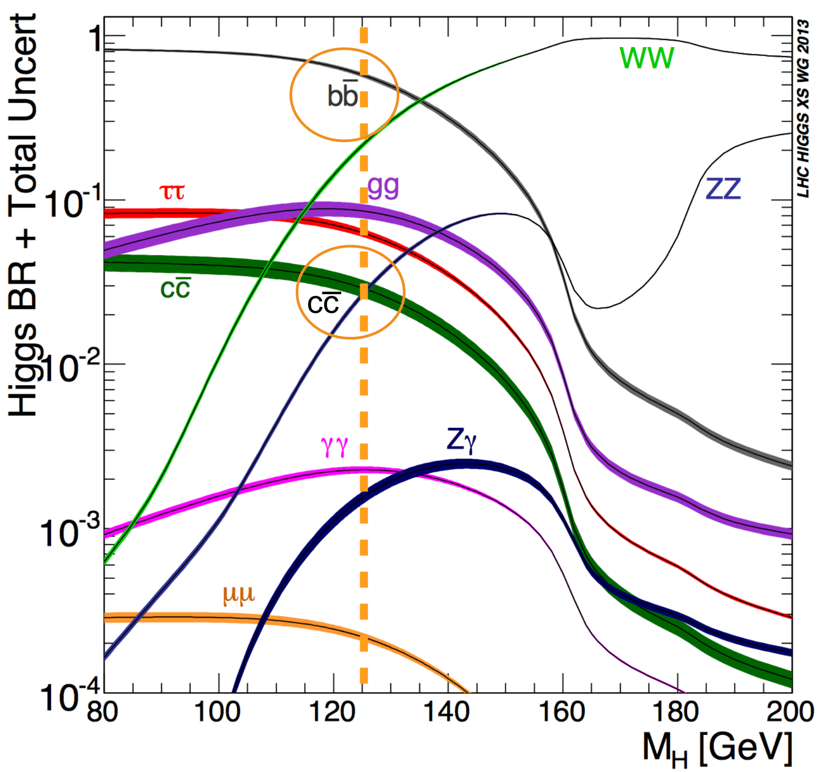
- **Introduction**
- **Higgs-b-quark coupling**
  - VH, H➔b$\bar{b}$ (2017 dataset)
  - Combination of H➔b$\bar{b}$ analyses
  - H➔b$\bar{b}$ prospects at HL-LHC
- **Higgs-c-quark coupling**
  - Differential distributions
  - Rare decays
- **Summary & outlook**

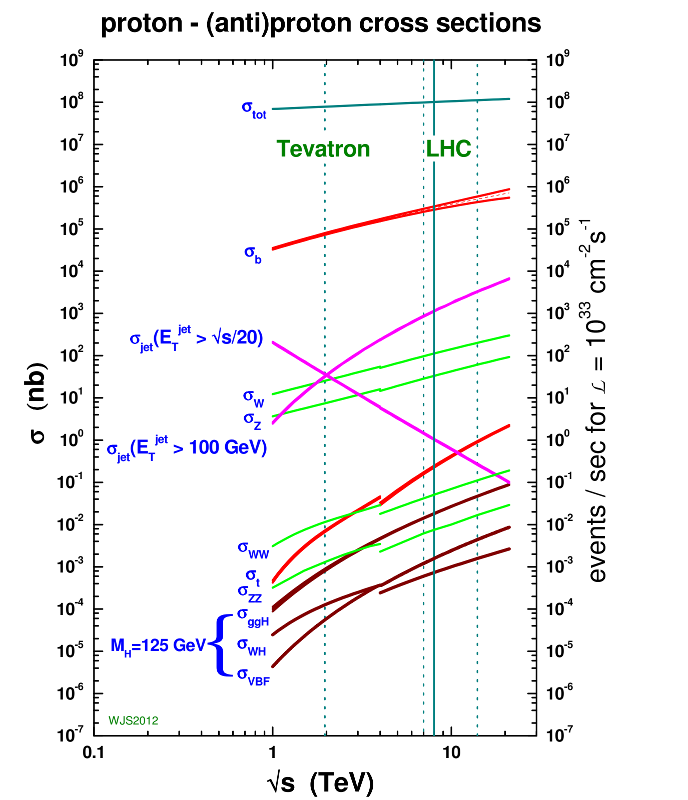
Why H→bb̄ and H→cc̄?

- **Largest Higgs boson branching ratio is into bb̄: 58%**
  - Largest contribution to total Higgs boson width
  - Access to down-type quark Yukawa coupling
- **Higgs boson branching ratio into cc̄: ~3%**
  - Access to 2nd generation quark coupling

Challenge: large background from multijet events

Challenge: as above + c-tagging



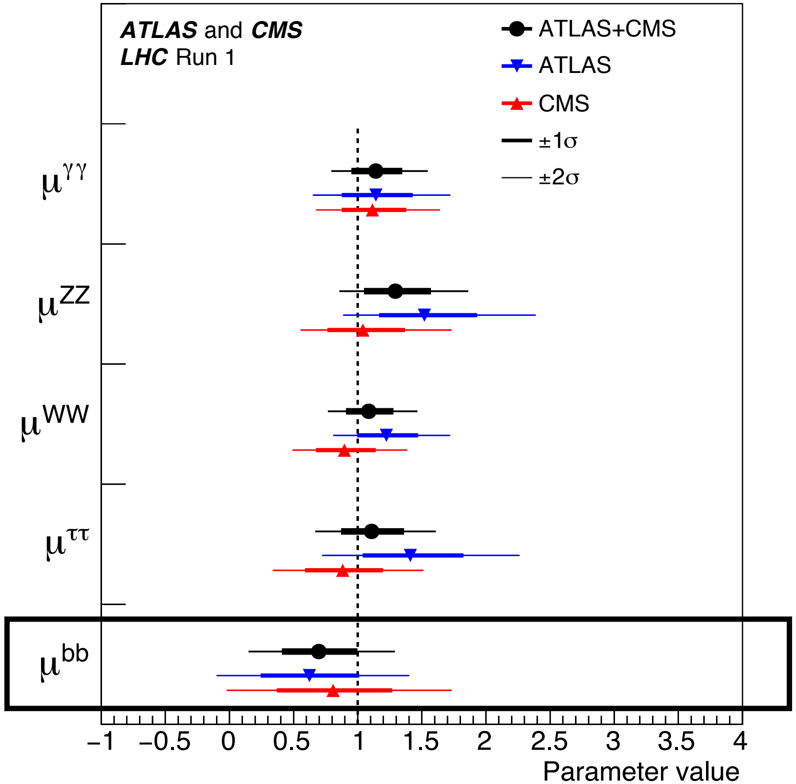


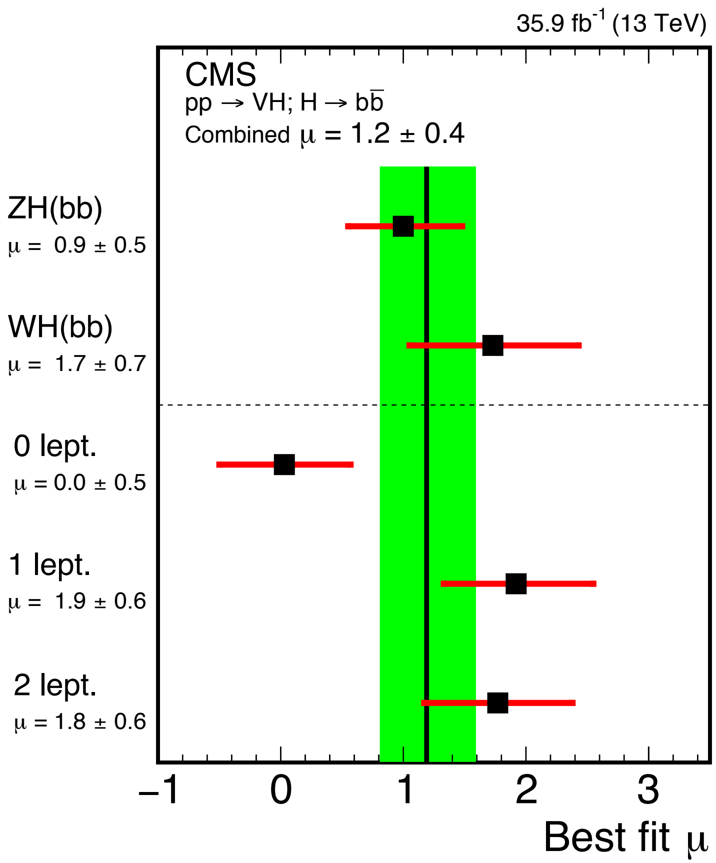
JHEP08(2016)045

PLB 780 (2018) 501-532

Previous results

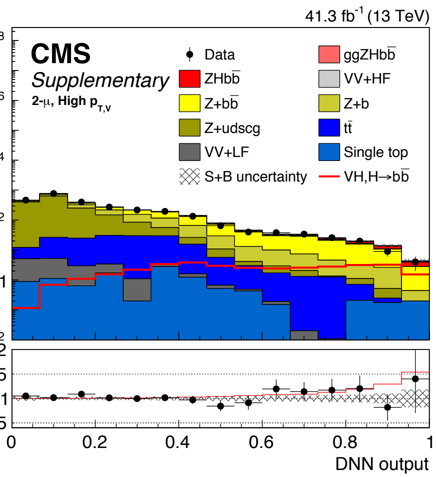
- Combining 7+8 TeV ATLAS and CMS results yielded a best-fit $\mu_{bb}=\sigma^{*}BR/(\sigma_{SM}^{*}BR_{SM}) = 0.7^{+0.29}_{-0.27}$
- CMS VH($b\bar{b}$) analysis at 13 TeV using the 2016 dataset: $\mu=1.19^{+0.21}_{-0.20}$ (stat) $^{+0.34}_{-0.32}$ (syst)
- Combined with CMS Run-1 VH($b\bar{b}$): $\mu=1.06^{+0.31}_{-0.29}$ (stat+syst): **Evidence for H→$b\bar{b}$**

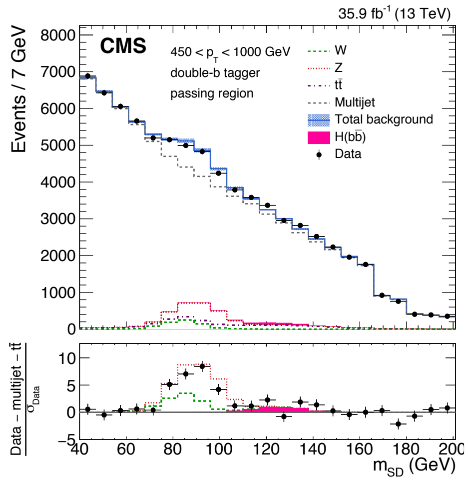


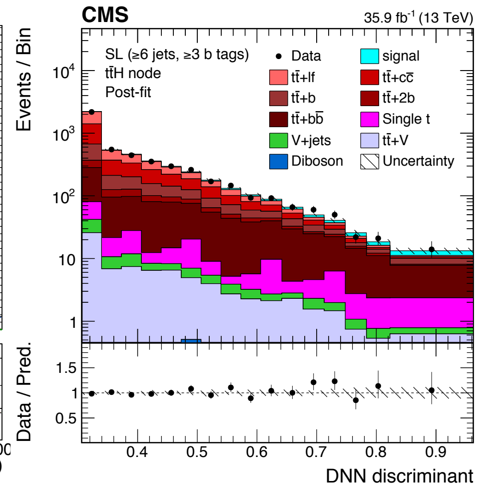


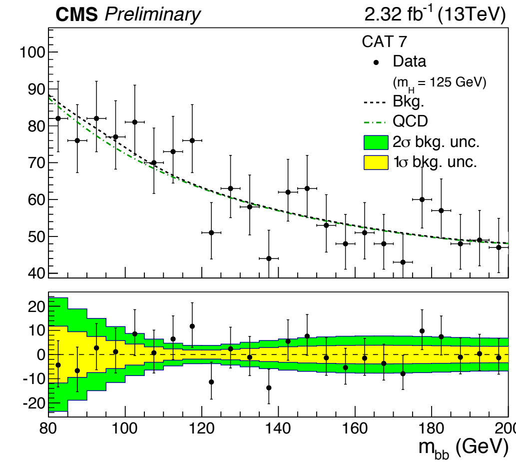
Higgs-b-quark coupling

$H\rightarrow b\bar{b}$ at CMS in Run-2









VH, $H\rightarrow b\bar{b}$

2016-dataset: 35.9 fb^{-1}
[PLB 780 (2018) 501-532]
2017-dataset: 41.3 fb^{-1}
[PRL 121(2018) 121801]

ggH, $H\rightarrow b\bar{b}$

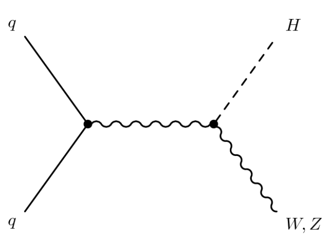
2016-dataset: 35.9 fb^{-1}
[PRL 120(2018) 071802]

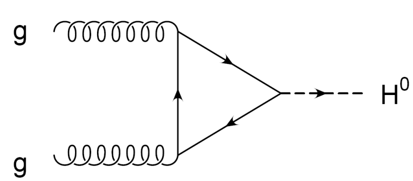
ttH, $H\rightarrow b\bar{b}$

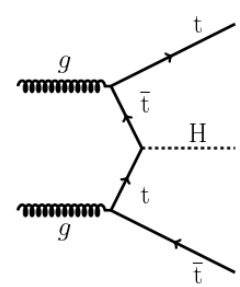
2016-dataset: 35.9 fb^{-1}
[JHEP 06 (2018) 101 & arXiv:1804.03682 (sub'd to JHEP)]

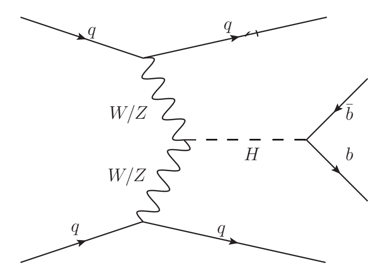
VBF $H\rightarrow b\bar{b}$

2015-dataset: 2.3 fb^{-1}
[CMS-PAS-HIG-16-003]





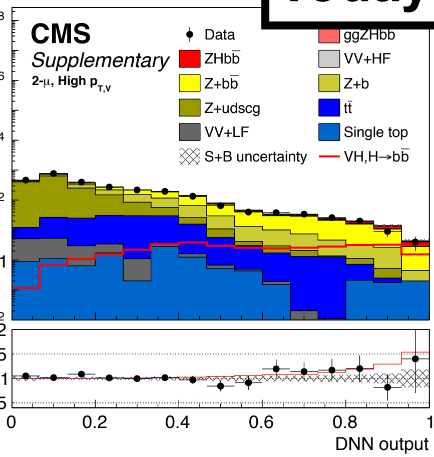


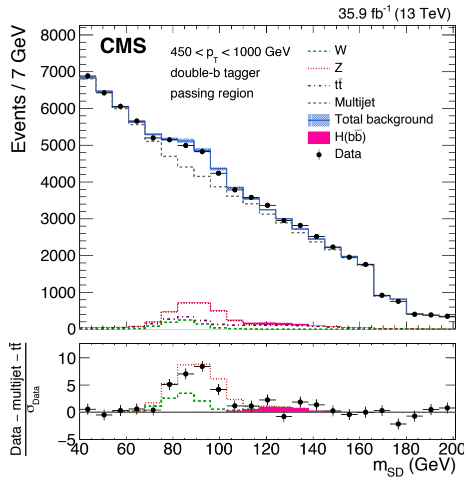


All major production modes covered in the $H\rightarrow b\bar{b}$ decay channel at CMS!

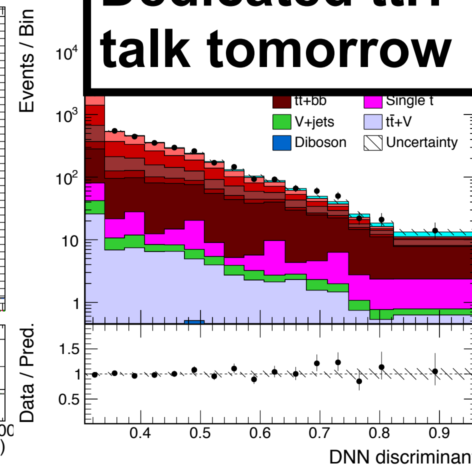
H→b$\bar{b}$ at CMS in Run-2

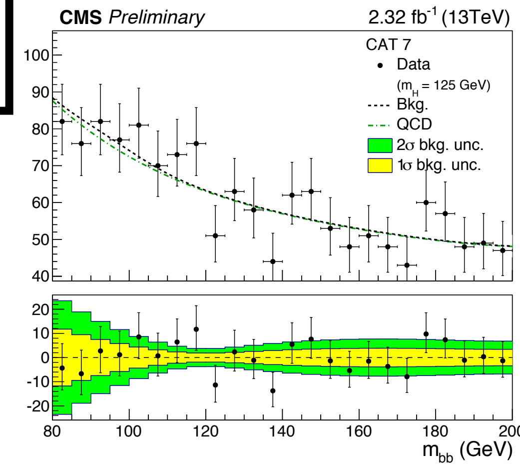
Today





Dedicated ttH talk tomorrow





VH, H→b$\bar{b}$

2016-dataset: 35.9 fb^{-1}
[PLB 780 (2018) 501-532]
2017-dataset: 41.3 fb^{-1}
[PRL 121(2018) 121801]

ggH, H→b$\bar{b}$

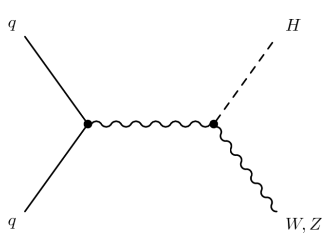
2016-dataset: 35.9 fb^{-1}
[PRL 120(2018) 071802]

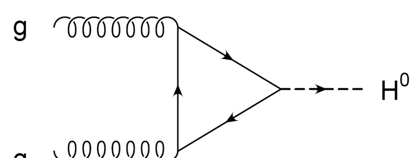
ttH, H→b$\bar{b}$

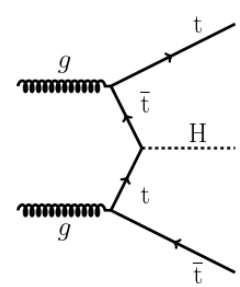
2016-dataset: 35.9 fb^{-1}
[JHEP 06 (2018) 101 & arXiv:1804.03682 (sub'd to JHEP)]

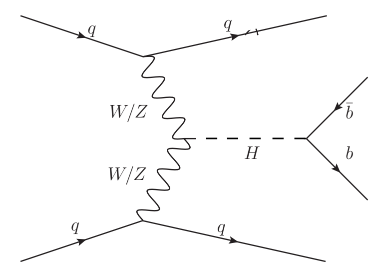
VBF H→b$\bar{b}$

2015-dataset: 2.3 fb^{-1}
[CMS-PAS-HIG-16-003]





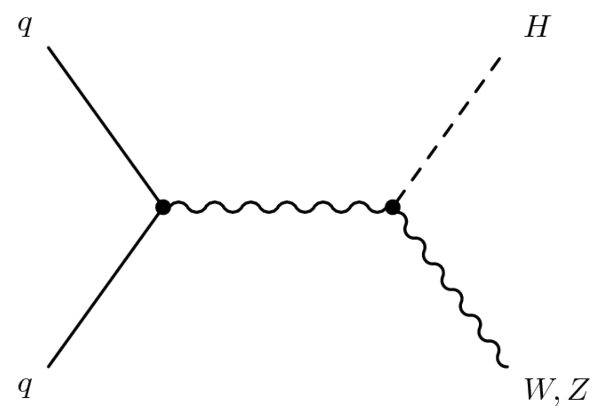




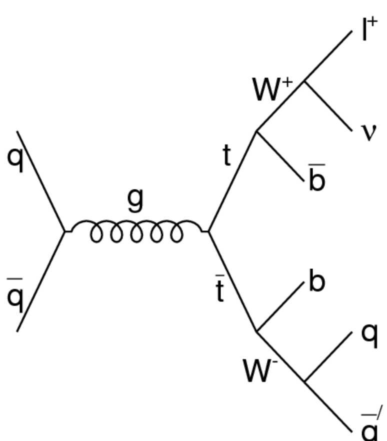
All major production modes covered in the H→b$\bar{b}$ decay channel at CMS!

VH, H➔b$\bar{b}$

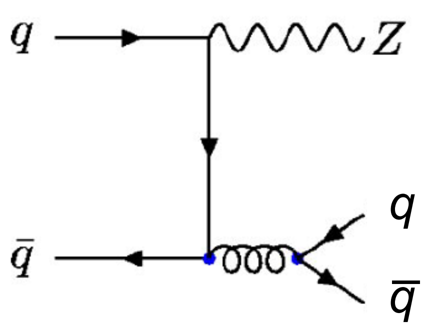
Overview



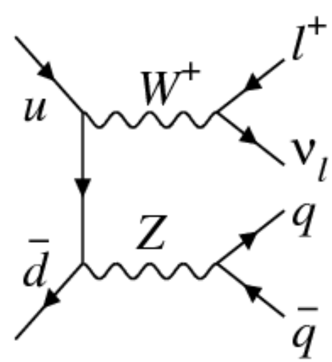
- Higgs boson produced in association with a vector boson
  - Leptonically decaying vector boson gives a clean signature to tag ➔ helpful for online selection
  - Much reduced background from multijet production
- Most sensitive channel for H➔b$\bar{b}$ despite relatively small production cross section
- **Backgrounds:**



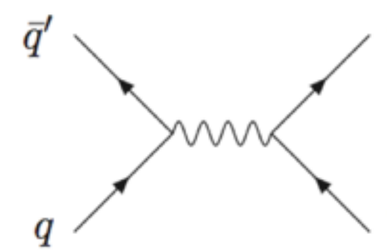
t$\bar{t}$



W,Z + jets



di-boson

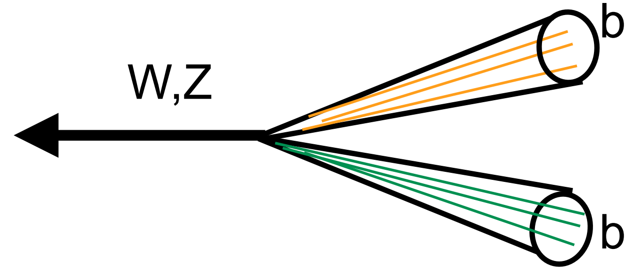


single top

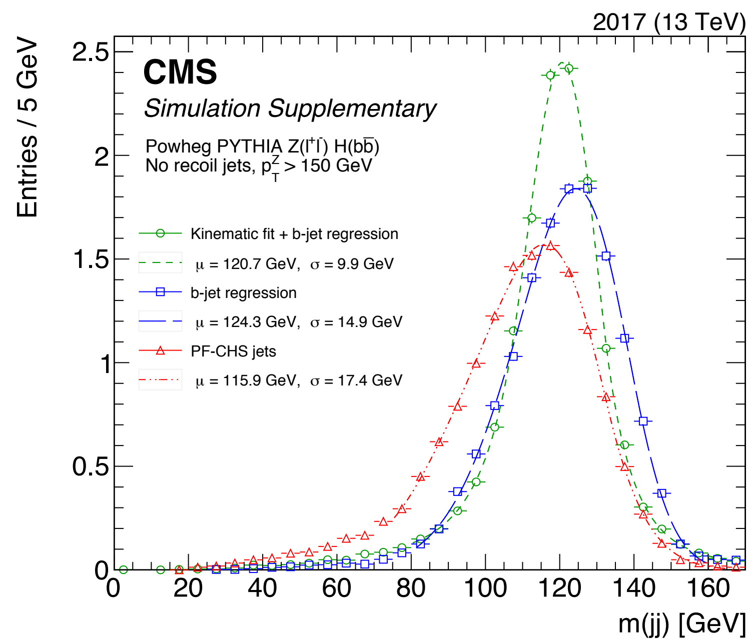
VH, H→$b\bar{b}$

Analysis strategy

- **Select** events with 0, 1, or 2 leptons (e/μ), consistent with W/Z decay, and 2 b-tagged jets
  - b-jets and vector boson produced back-to-back + increased sensitivity for enhanced Higgs boson p_T ➔ **categorise** based on V p_T
- **Improve the m_{bb} resolution**
  - using multivariate regression techniques
  - using kinematic fit in 2-lepton channel
- Use a Deep Neural Network (DNN) to increase the separation between signal and background
  - Variables include V p_T, m_{bb}, jet kinematics, b-tagging information, ...
- **Fit for signal** using the DNN as final discriminant, simultaneously fitting control regions to constrain the major backgrounds
- **Validate** the analysis strategy using a di-boson analysis as well as a fit of the di-b mass distribution.

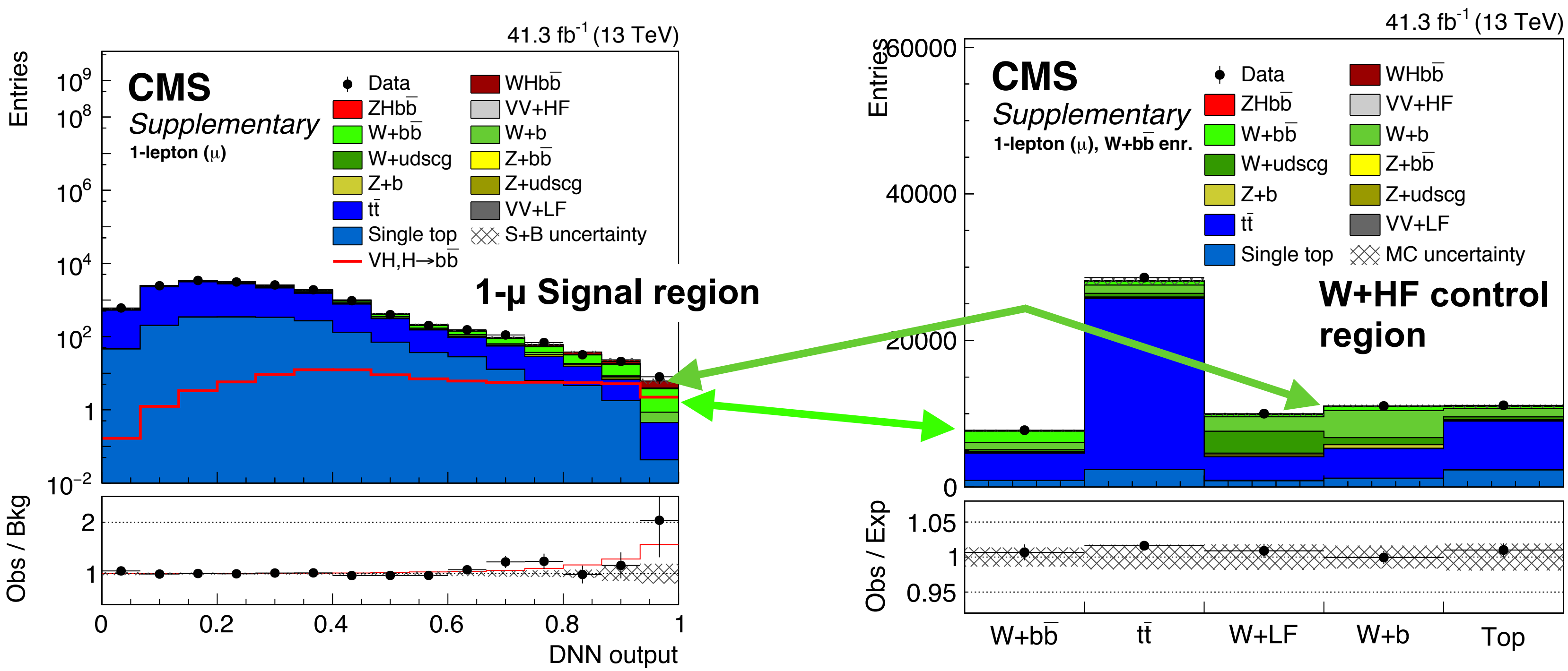


	0-lepton	1-lepton	2-lepton	
V p_T	> 170 GeV	> 150 GeV	50-150 GeV	> 150 GeV



Background normalisation

- **Example: constraining background normalisations in 1-lepton channel**
  - **$t\bar{t}$** and **W+LF** background: single-bin control regions
  - **W+HF** background: DNN multi-classifier to distinguish between background components
- Similar strategy in 0-lepton channel, 2-lepton channel Z+HF control region is more pure and we fit the b-tagging discriminator distribution in 2 bins.



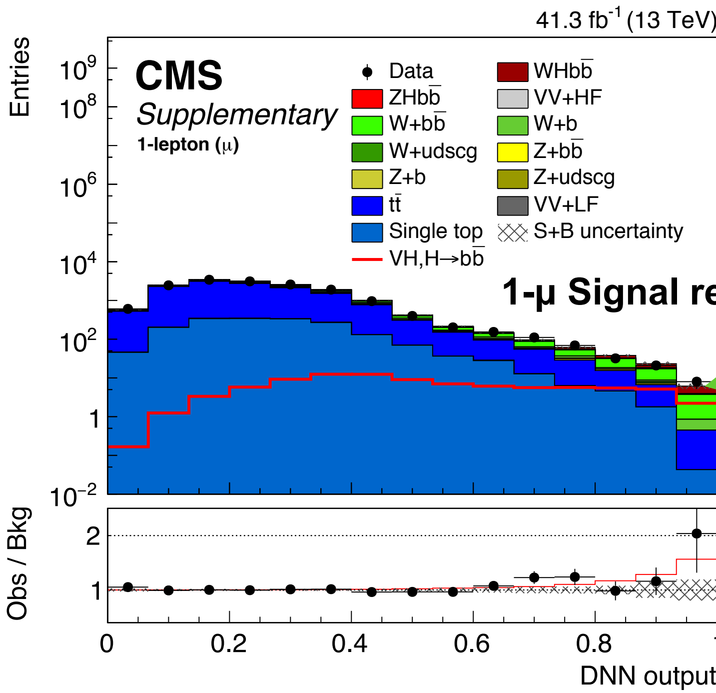
Background normalisation

- **Example: constraining background normalisations in 1-lepton channel**
  - **$t\bar{t}$** and **W+LF** background: single-bin control regions
  - **W+HF** background: DNN multi-classifier to distinguish between background components

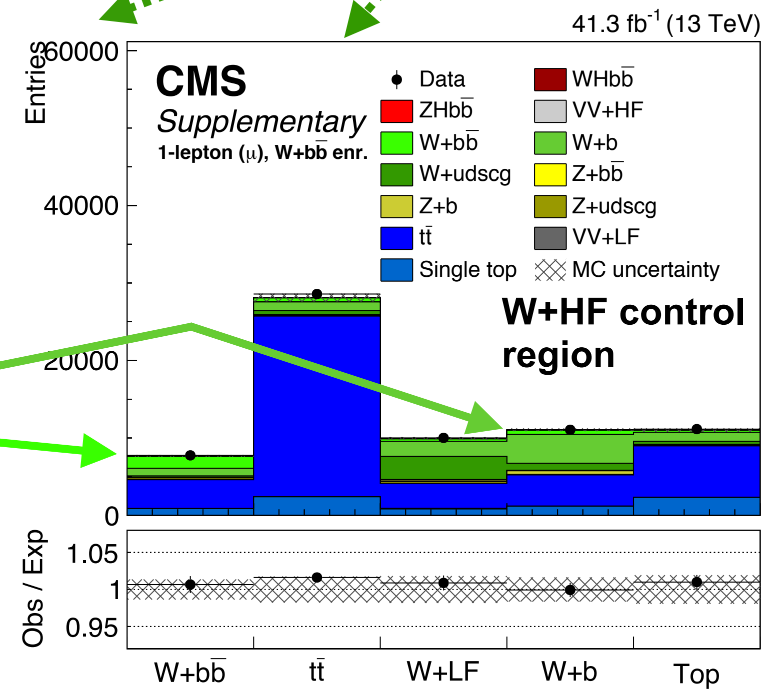
- Similar strategy in 0-lepton channel, 2-lepton channel Z+HF control region is more pure and we fit the b-tagging discriminator distribution in 2 bins.

+Normalisation in $t\bar{t}$ enriched CR

+Normalisation in W+LF enriched CR



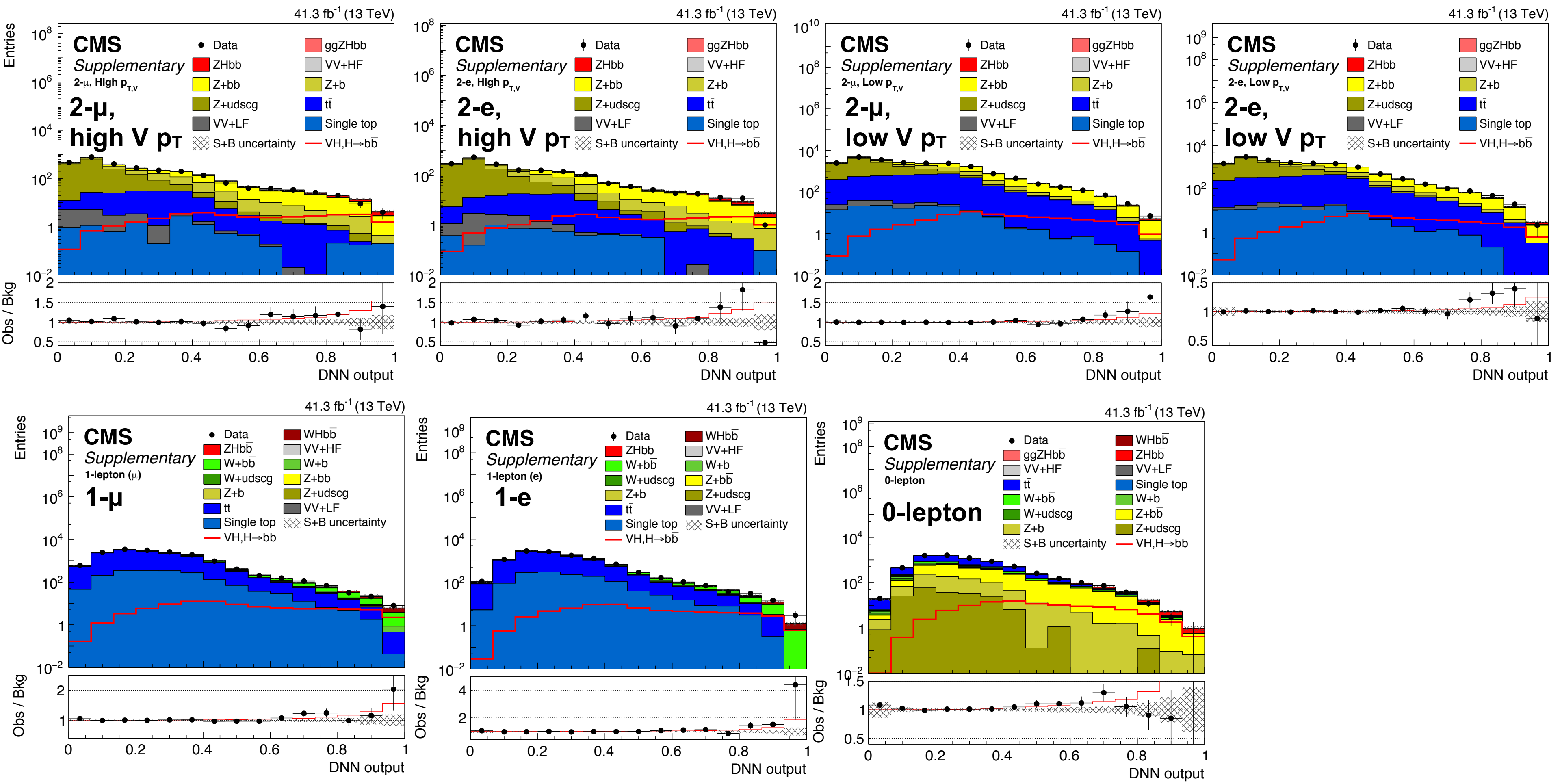
1-μ Signal region



W+HF control region

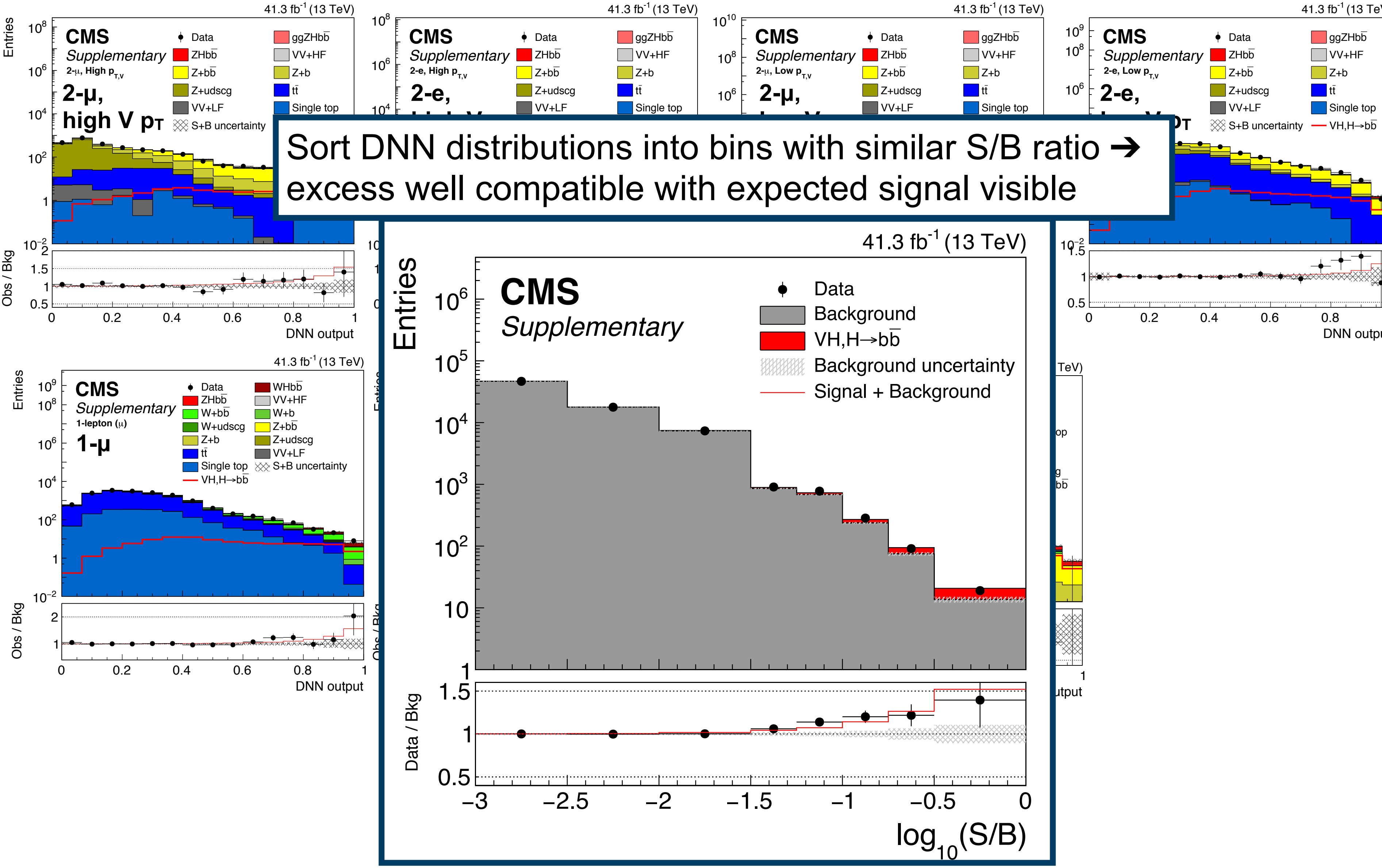
PRL 121(2018) 121801

A closer look at the signal regions



PRL 121(2018) 121801

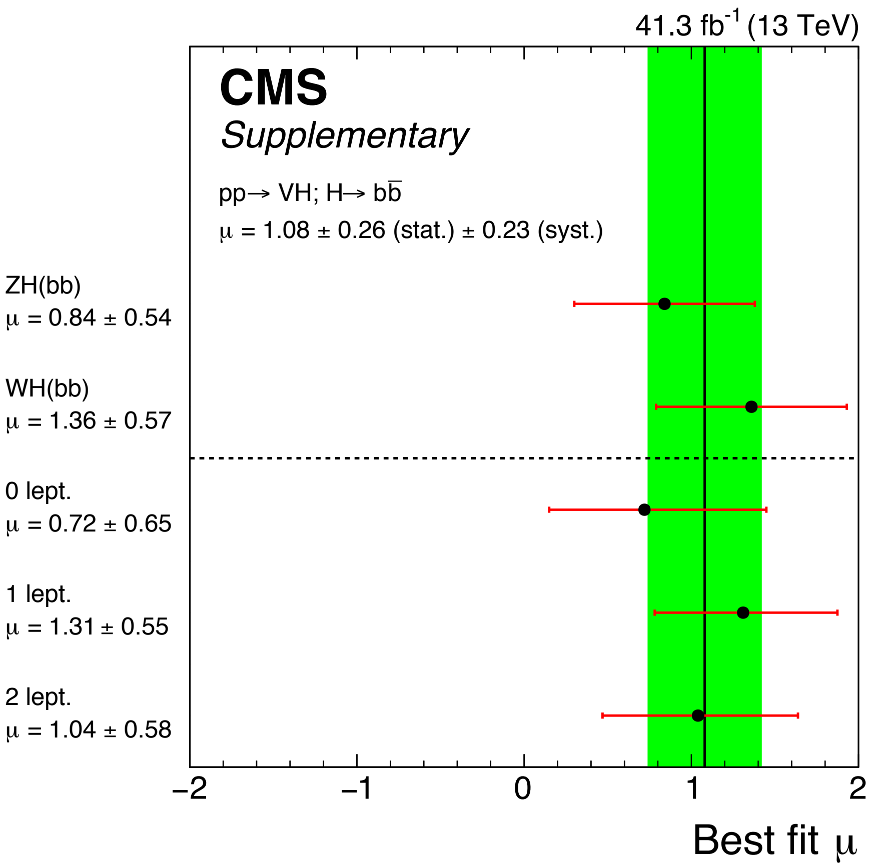
A closer look at the signal regions



Sort DNN distributions into bins with similar S/B ratio ➔ excess well compatible with expected signal visible

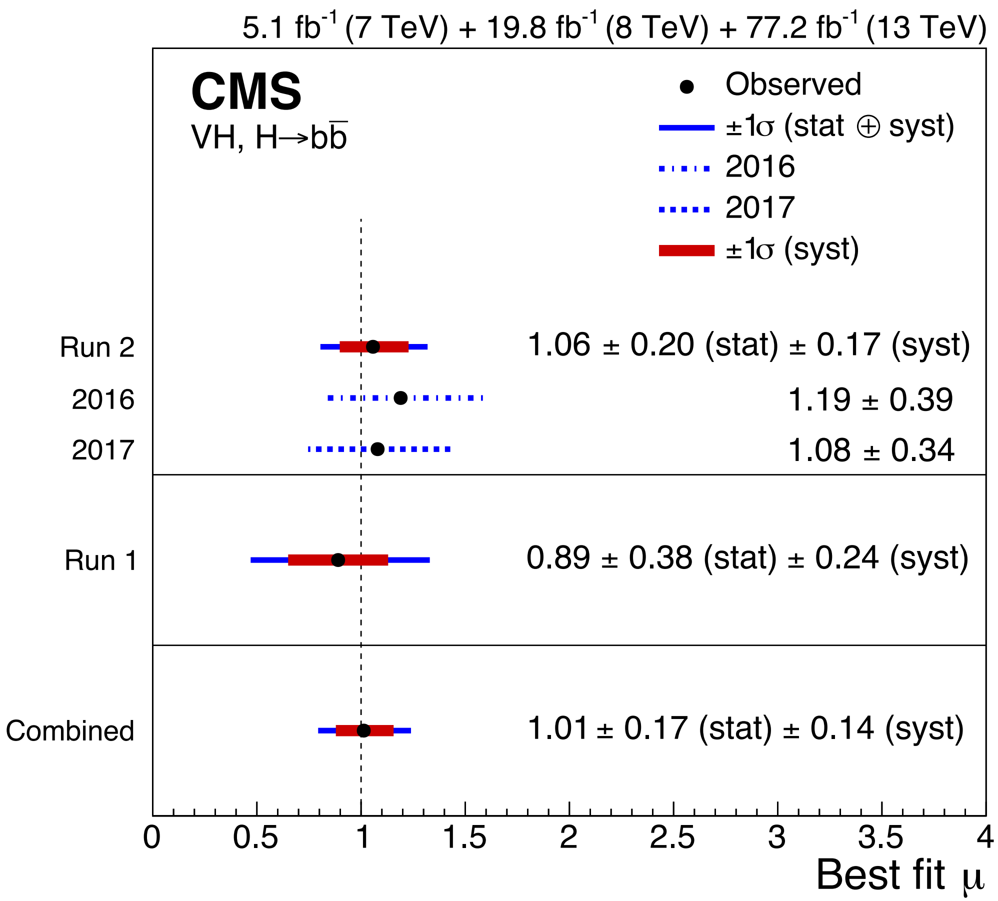
Results

- **Result of 2017 analysis: μ=1.08 ± 0.26 (stat.) ± 0.23 (syst.)**
  - Compatible with SM expectation
- **Total uncertainty ~0.34, statistical and systematic component of similar order**
  - Major sources of systematic uncertainty: background normalisation, size of simulated samples, b-tagging



Uncertainty source	$\Delta\mu$	
Statistical	+0.26	−0.26
Normalization of backgrounds	+0.12	−0.12
Experimental	+0.16	−0.15
b-tagging efficiency and misid	+0.09	−0.08
V+jets modeling	+0.08	−0.07
Jet energy scale and resolution	+0.05	−0.05
Lepton identification	+0.02	−0.01
Luminosity	+0.03	−0.03
Other experimental uncertainties	+0.06	−0.05
MC sample size	+0.12	−0.12
Theory	+0.11	−0.09
Background modeling	+0.08	−0.08
Signal modeling	+0.07	−0.04
Total	+0.35	−0.33

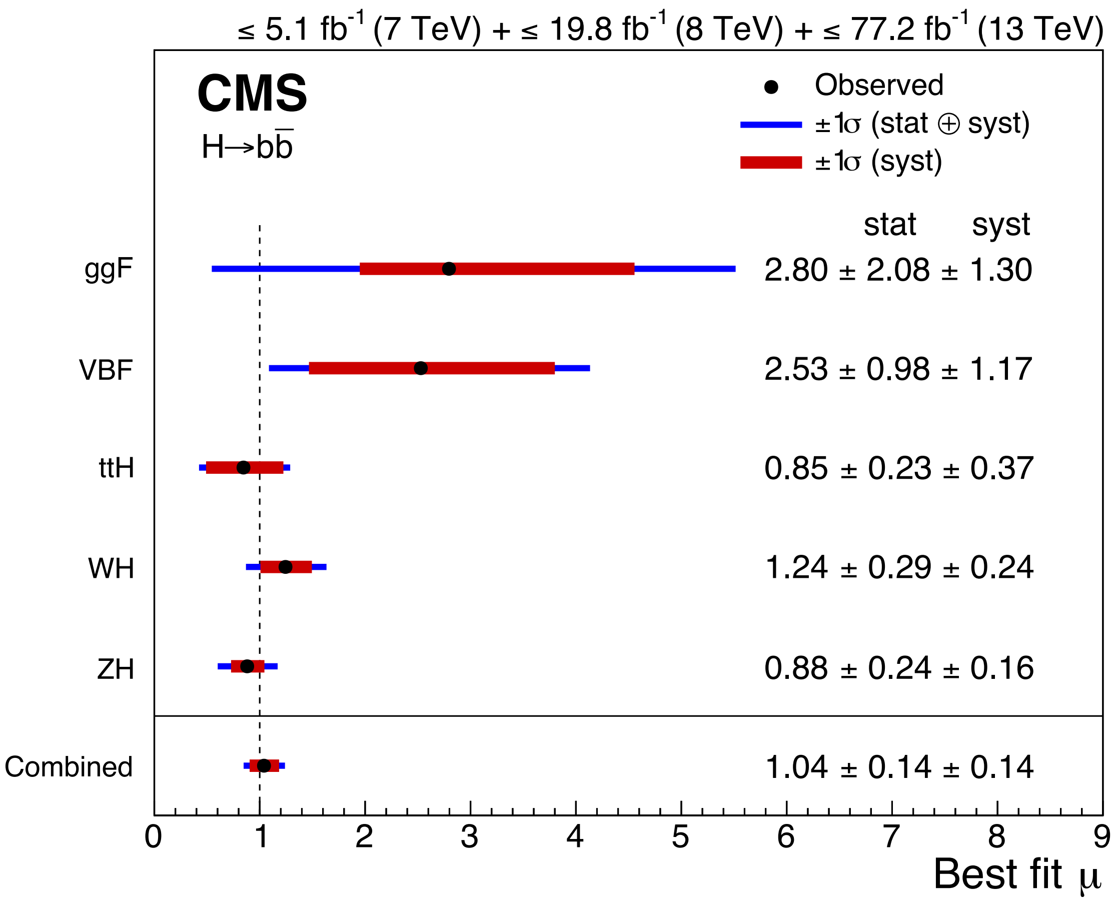
Combination of VH, H→$b\bar{b}$ analyses



- Combining with VH($b\bar{b}$) analysis on 2016 dataset and Run 1 VH($b\bar{b}$):
  - **μ=1.01±0.17 (stat.) ± 0.14 (syst.)**
  - Significance 4.8σ (4.9σ) obs (exp)

PRL 121(2018) 121801

Combination of H→b$\bar{b}$ analyses

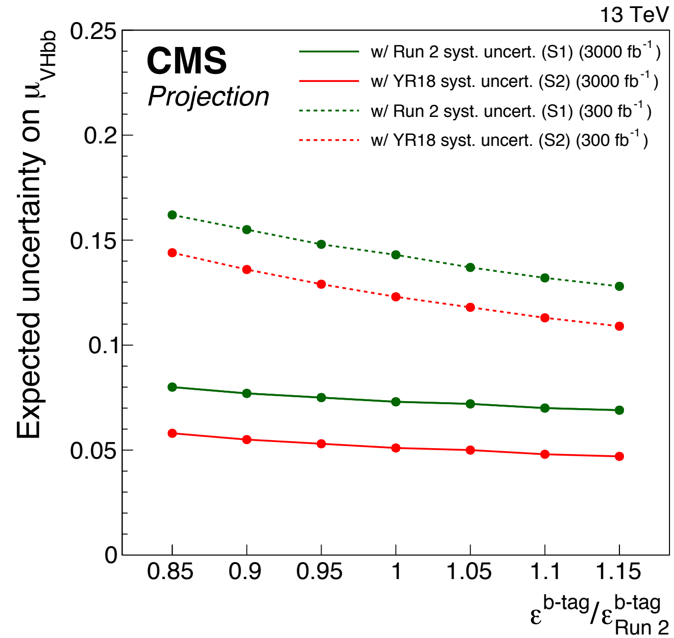


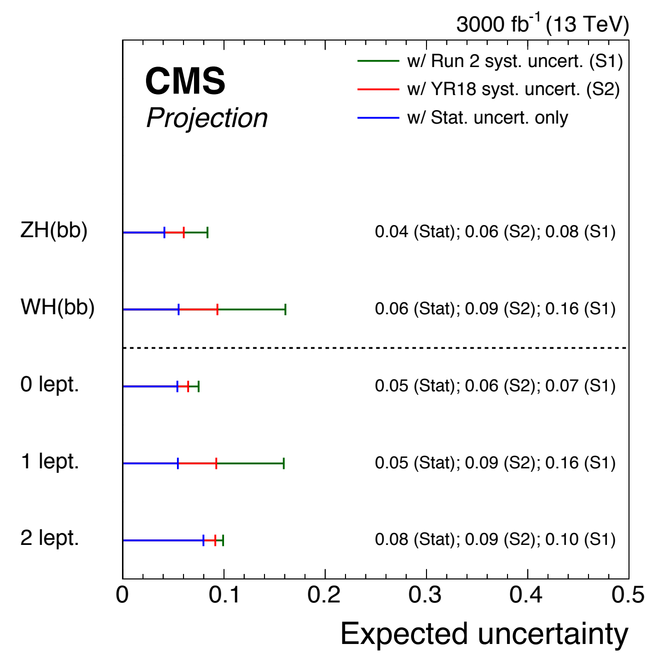
- Combine VH(b$\bar{b}$) results with:
  - Run 1 + 2016 ttH(b$\bar{b}$)
  - 2016 boosted ggH(b$\bar{b}$)
  - Run 1 VBF (b$\bar{b}$)
- **μ=1.04 ± 0.14 (stat) ± 0.14 (syst)**
- observed (expected) significance 5.6σ (5.5σ)

VH, H→b$\bar{b}$ at HL-LHC

- Consider various scenarios for uncertainties:
  - **With Run-2 systematic uncertainties**: uncertainties as in Run 2 (S1)
  - **With YR18 systematic uncertainties**: most experimental uncertainties scale down with sqrt(L), until a lower limit is reached. Theoretical uncertainties are ~halved. (S2)
  - **Stat. Only.:** No systematic uncertainties considered
- At 3 ab^{-1}, measurement will be driven by theoretical uncertainties, ggZH QCD scale uncertainty becomes important.
- All channels contribute ~equally: challenge experimentally to maintain trigger thresholds.
- Effect of changing b-tagging efficiency is non-negligible.

	S1	S2
Total uncertainty	7.3%	5.1%
Signal theory uncertainty	5.4%	2.6%
Inclusive	4.6%	2.2%
Acceptance	2.7%	1.3%
Background theory uncertainty	2.8%	2.3%
Experimental uncertainty	2.6%	2.2%
b-tagging	2.2%	2.0%
JES and JER	0.7%	0.6%
Statistical uncertainty	3.2%	3.2%

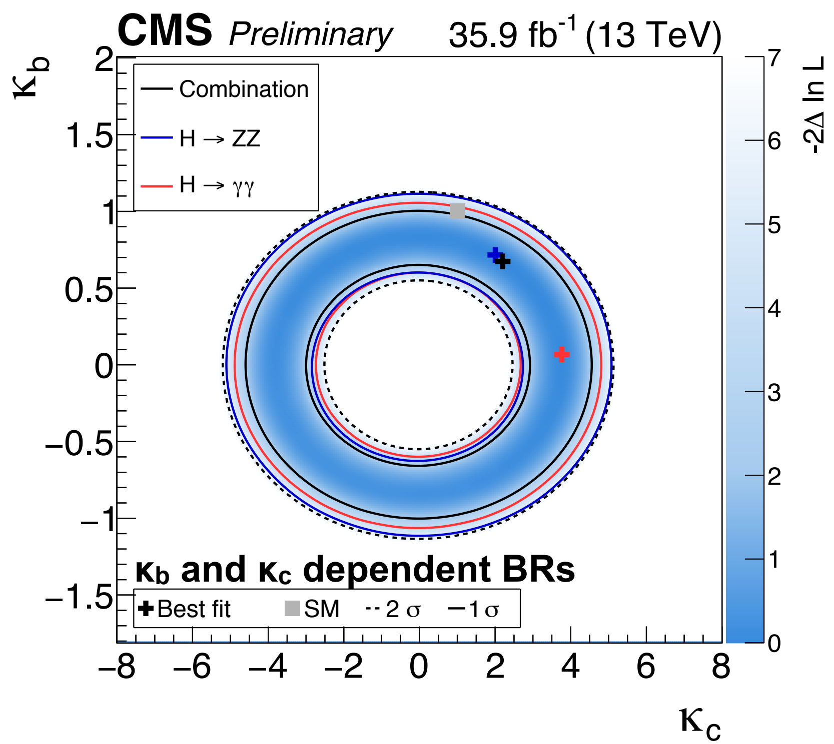


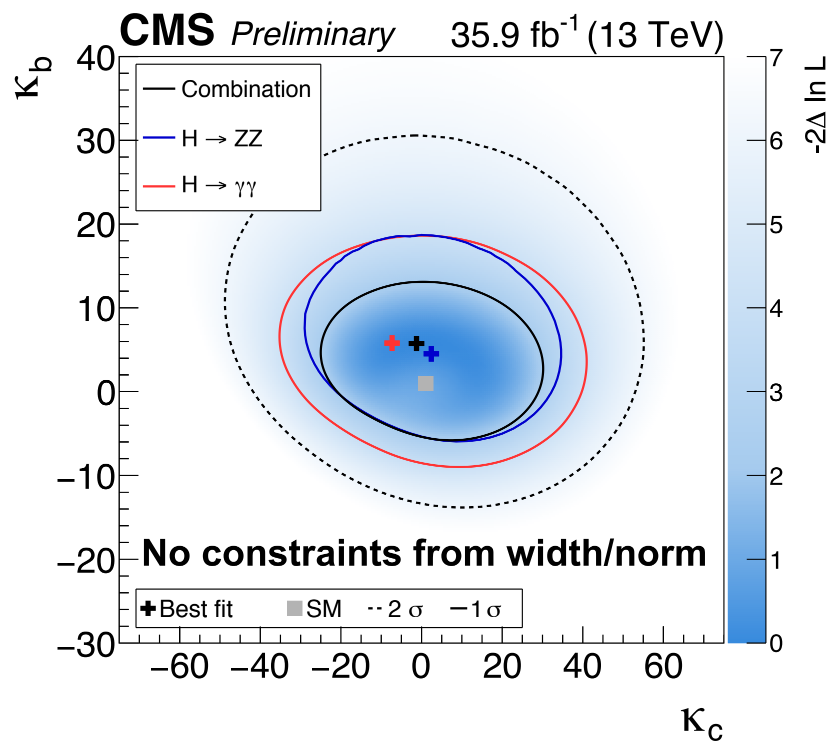


Higgs-c-quark coupling

Higgs-charm coupling from differential distributions

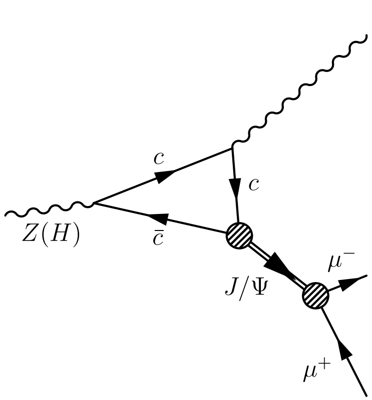
- Using the κ-model to describe effective coupling between Higgs boson and other particles, the $p_T(H)$ spectrum is sensitive to modifications in κ_b and κ_c.
  - Apply this to $p_T(H)$ differential measurements in H→γγ and H→ZZ
- Two assumptions tested:
  - Branching ratios depend on κ_b and κ_c : **$-4.3 < \kappa_c < 4.3$ (exp: $-5.4 < \kappa_c < 5.3$)** (±1σ interval)
  - Branching ratios freely floating (no constraints from total Higgs width and overall normalisation): **$-18.0 < \kappa_c < 22.9$ (exp: $-15.7 < \kappa_c < 19.3$)** (±1σ interval)



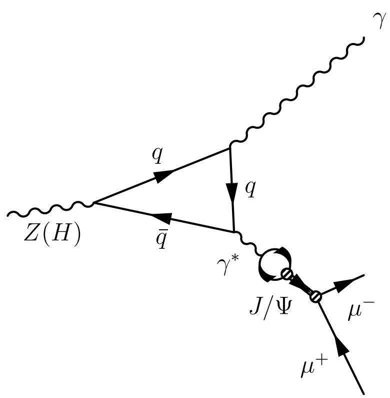


Higgs-charm coupling from rare Higgs decays

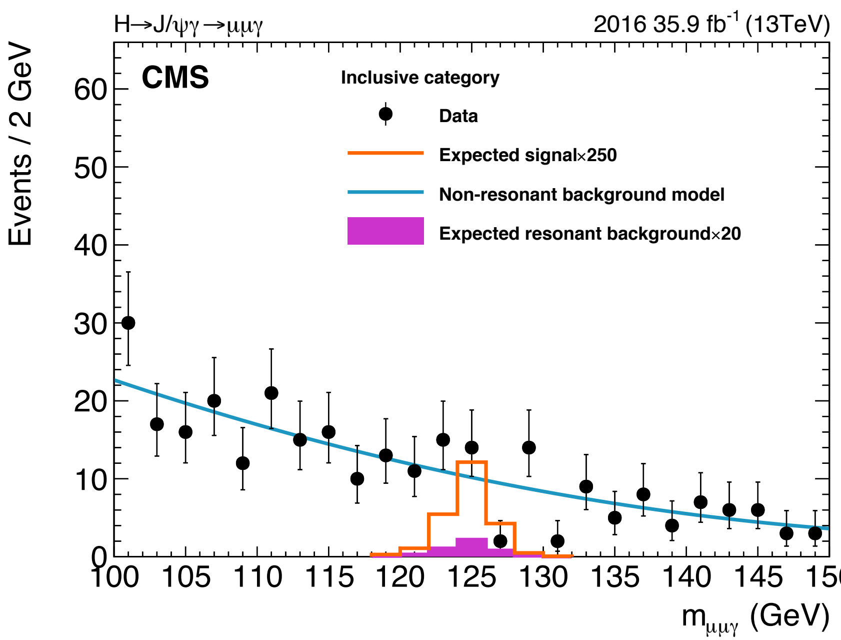
- Higgs boson can decay to γ and a $c\bar{c}$ resonance (e.g. J/ψ) ➔ small SM branching ratio ($3 \cdot 10^{-6}$)
- Decay does not always include the H-c vertex.
- Search for H➔γJ/ψ and Z➔γJ/ψ, performed using 35.9 fb^{-1} collected in 2016
- Use μμγ invariant mass as discriminating variable in both cases, with different signal mass regions to target H or Z decay.
- H➔γJ/ψ results:
  - Observed (expected) limit 260 x SM (170 x SM)
  - Combining with 8 TeV analysis: 220 x SM (160 x SM)



Direct process



Example of indirect process



Summary and outlook

- Presented $H \rightarrow b\bar{b}$ results and constraints on Higgs-c coupling from CMS

- $H \rightarrow bb$ observation in CMS by combination of analyses targeting all major Higgs boson production modes, sensitivity driven by VH and $t\bar{t}H$
  - **We are entering the precision era → shift focus towards more differential measurements**

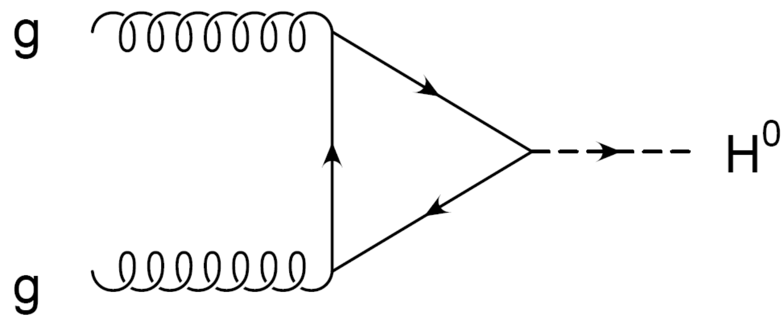
- Constraints on Higgs-c coupling from differential distributions and searches for $H \rightarrow \gamma J/\psi$
  - Direct $H \rightarrow c\bar{c}$ results coming

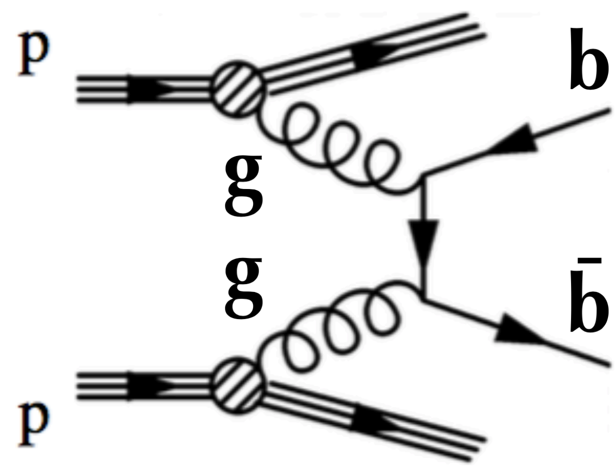
Backup

ggH, H→bb

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→bb 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→bb search at high H p_T possible**

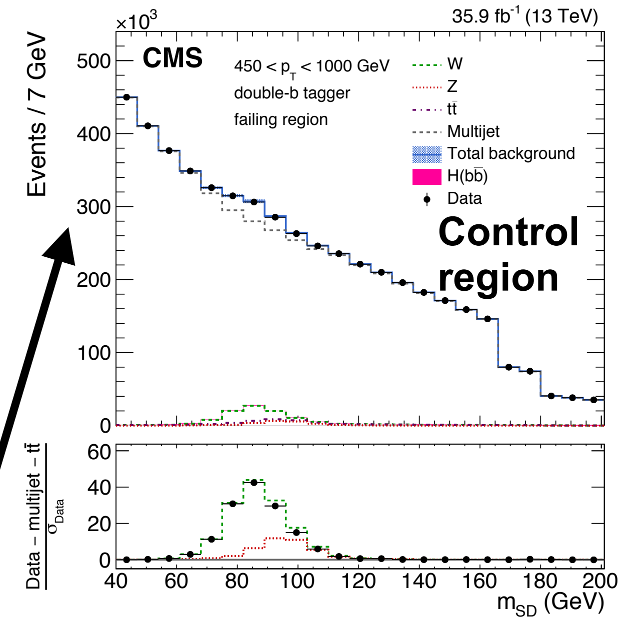


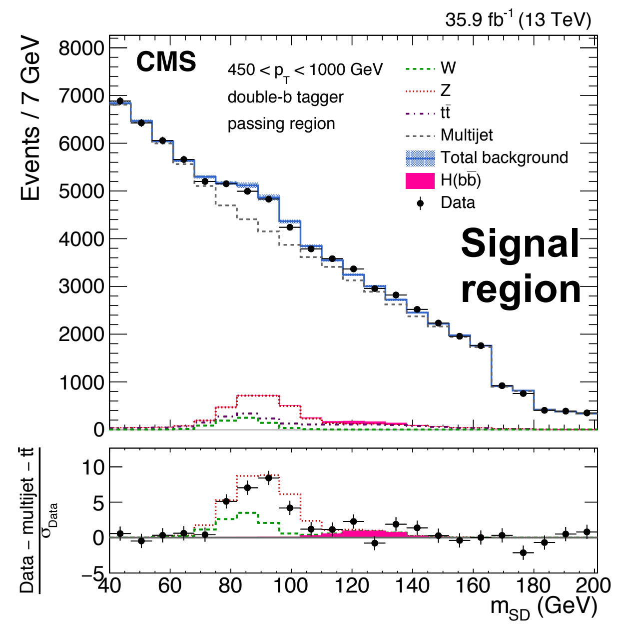


PRL 120(2018) 071802

Inclusive (ggH) H→bb

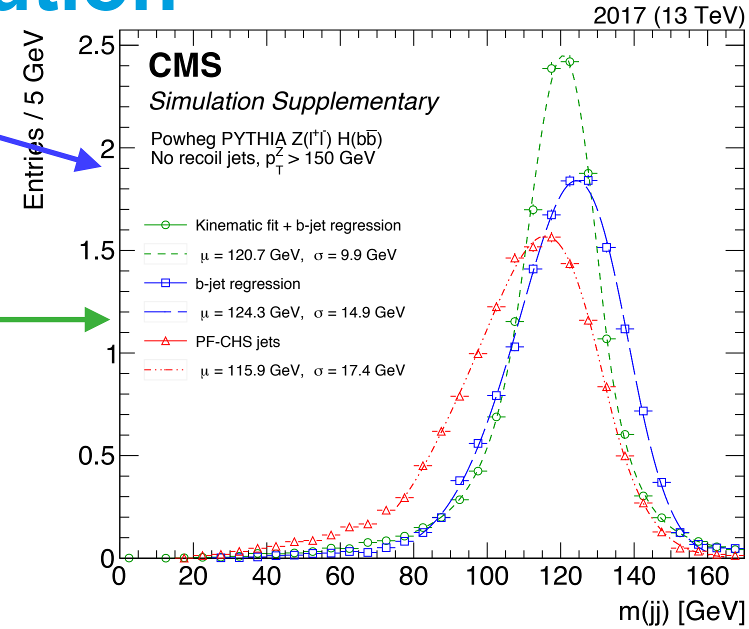
- Analysis performed using 2016 dataset
- Require a high p_T (>450 GeV) wide-cone jet and **exploit 2-prong jet substructure** and **b-tagging information of the subjects** to reduce multijet background
- Soft radiation removed from candidate jet to provide better separation between signal and multijet background when using the jet mass shape
- **Backgrounds** mostly from multijet production, with smaller contributions from tt, W and Z production
    - **Multijet** background estimated from a data sample with double b-tag requirement inverted. Simultaneous fit with SR
- Result: $\mu_H = 2.3 \pm 1.5$ (stat.) +1-0.4 (syst)



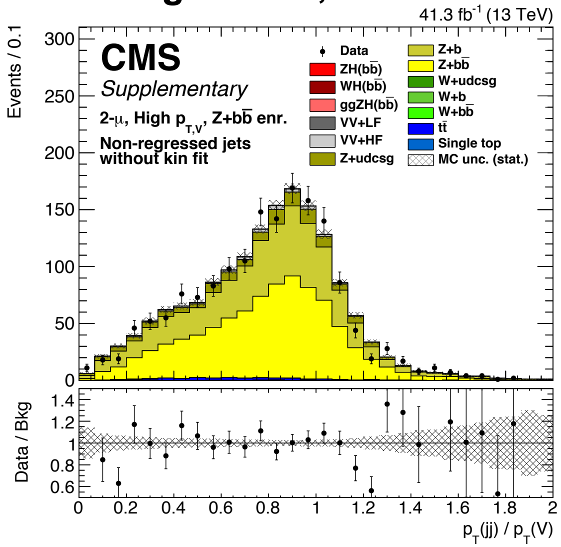


Improving di-b mass resolution

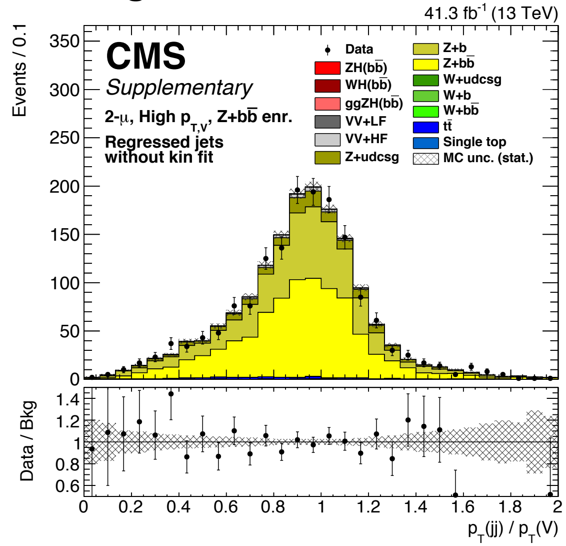
- In all channels, di-b mass resolution improved with DNN b-jet energy regression
- In 2-lepton channel, additional improvement from use of kinematic fit
- Validation using $p_T(jj)/p_T(ll)$ in 2-lepton Z+HF control region ➔ data described well by simulation after all techniques



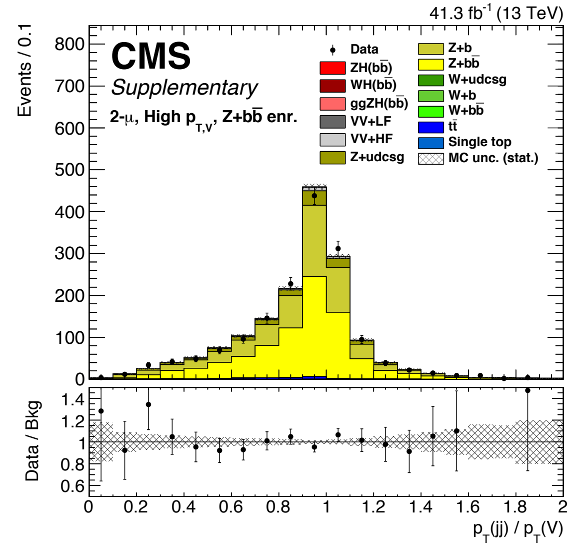
No regression, no kin fit



Regression, no kin fit



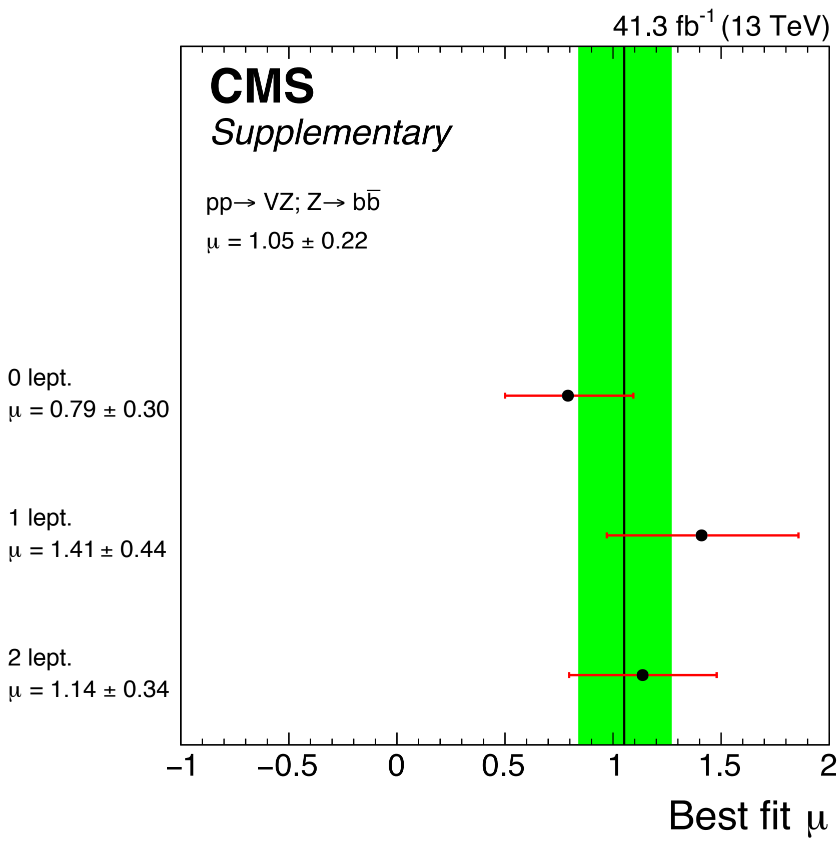
Regression & kin fit



VH, H→b$\bar{b}$ cross checks

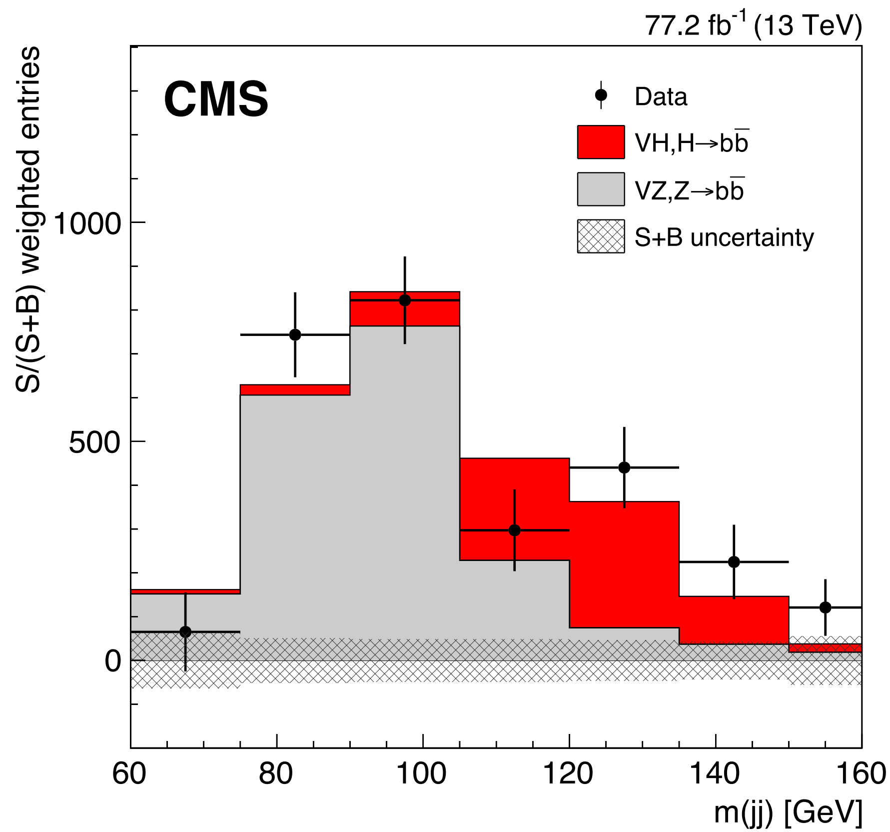
Extraction of VZ, Z→b$\bar{b}$

- Used to validate MVA methods: re-train DNN to separate VZ, Z→bb from other backgrounds, extract signal strength
- **μ=1.05±0.22: compatible with SM expectation**



Di-b mass analysis

- Fit the di-b mass distribution instead of DNN output
- Combination of analyses using 2016+2017 dataset, with backgrounds other than VZ subtracted



Talking points

- **Talking points:**
  - Interest from the theory community in measuring ggZH and qqZH separately (beyond STXS)?
  - In relation to the above point: are there possible solutions for the large ggZH QCD scale uncertainty?
  - Most interesting differential distributions for VH? (H $p_{T,}$ $m_{VH,}$ any others?)
  - Are there any other avenues that we are not currently exploring but should be?