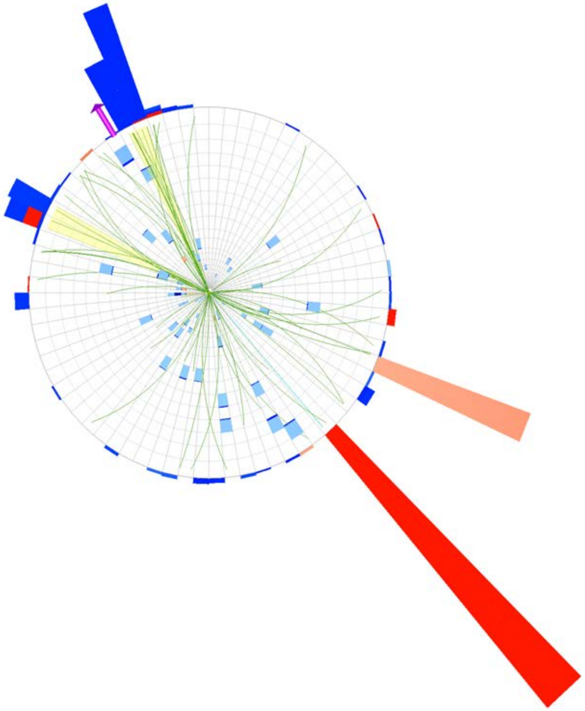
# Observation of the $H \rightarrow b\bar{b}$ decay at the CMS experiment



Rainer Mankel
DESY
for the CMS collaboration

*Terascale Alliance Annual Meeting*
*26-28 November 2018*
*Hamburg*

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

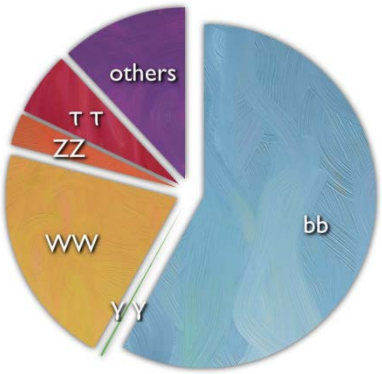
# Higgs bosons & Yukawa couplings

- Higgs Yukawa couplings to fermions are special
  - unlike any other experimentally probed interaction
  - ➔ a main goal of LHC Run-II

- H→bb is the dominant decay channel (58%)
  - probes coupling to a down-type quark
  - drives uncertainty of total Higgs boson width
  - important for constraining the BSM coupling
  - difficult to observed due to huge backgrounds
    - has been searched for at the LHC for the last six years
    - .. and earlier at LEP and Tevatron

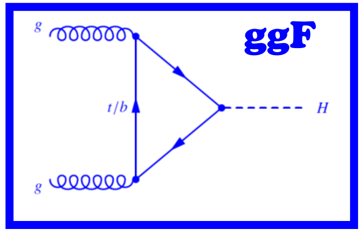
- Requires heavy use of machine learning techniques

| Higgs boson Yukawa couplings | | |
|---|---|---|
| *f* | Mass [GeV] | Observed |
| τ | 1.78 | August 2017 ☑ |
| t | 172.44 | April 2018 ☑ |
| **b** | **4.18** | **?** |

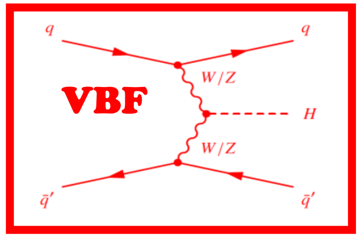


# Higgs boson production at the LHC

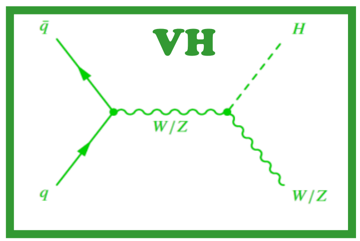
gluon fusion (~87%)



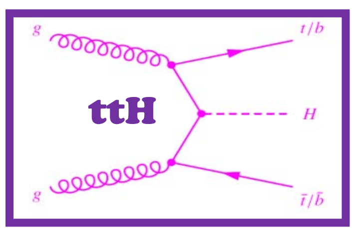

vector boson fusion (~7%)

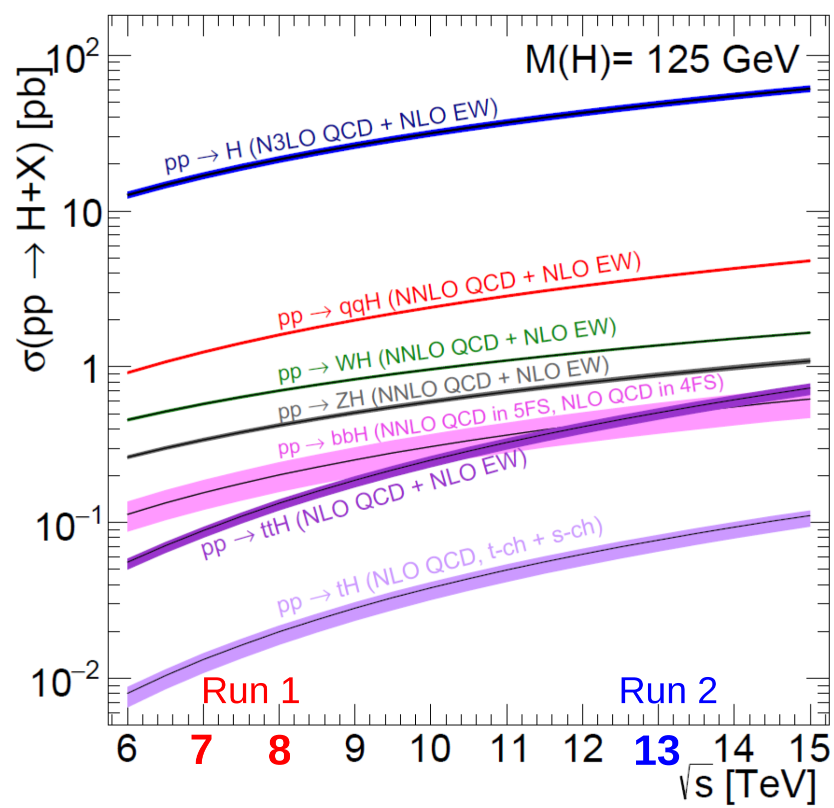



WH, ZH (~4%)




ttH (~1%)

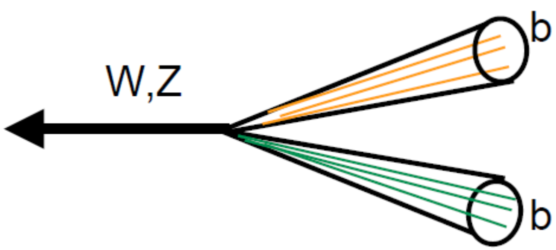


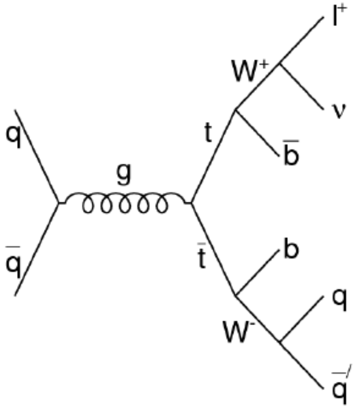

- ➔ VH production mode provides a clean signature for event selection and triggering
- ➔ highest sensitivity for H→bb

# Signature and backgrounds

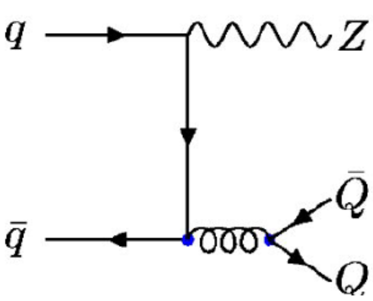
- Higgs boson produced in association with a vector boson (W or Z)
  - three channels: 0ℓ (Z→νν), 1ℓ (W→ℓν), 2ℓ (Z→ℓℓ)
  - good signature for triggering
  - strong reduction of multijet background



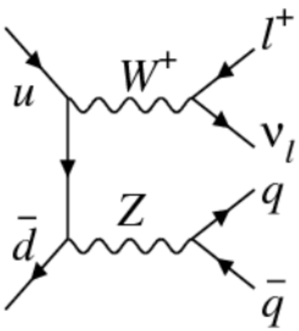

- Prominent backgrounds:



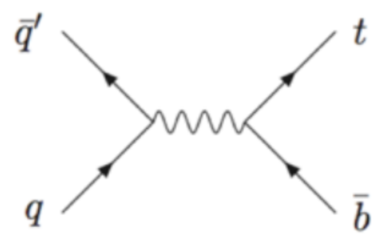

$t\bar{t}$ production




W, Z + jets




di-boson



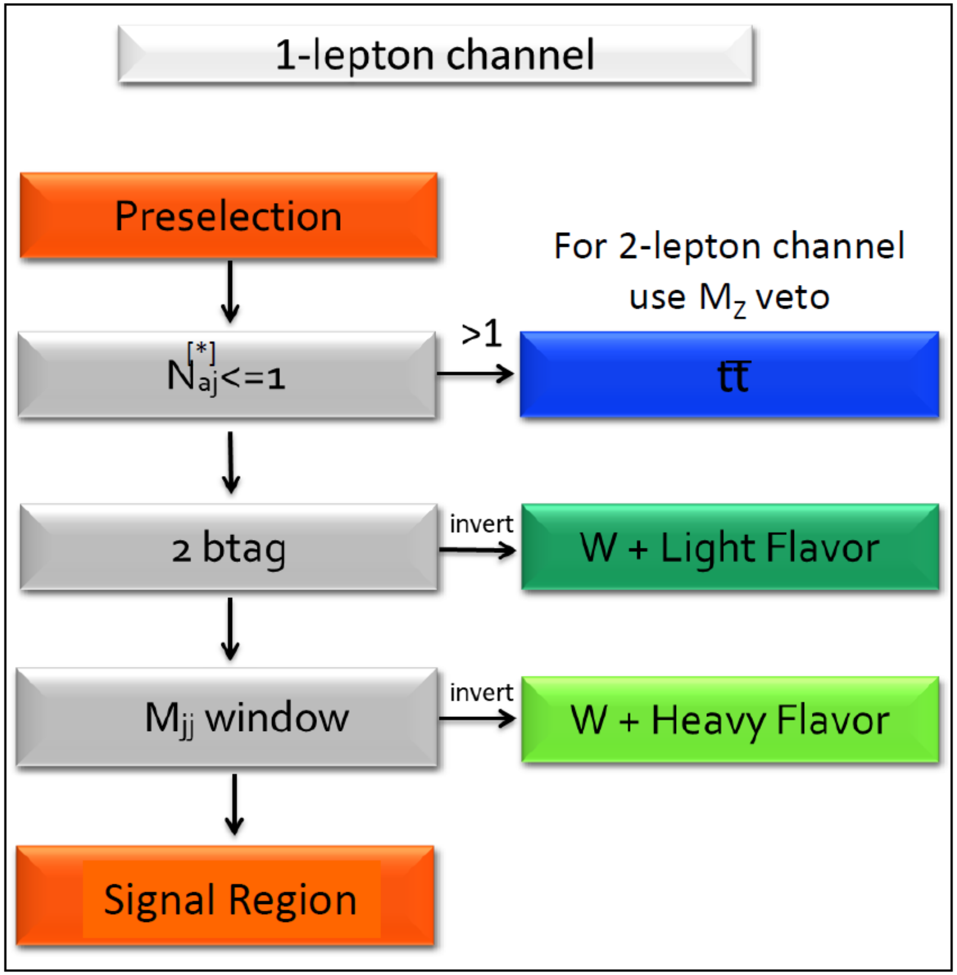

single-top

# Event selection & categorization

- Four analysis categories:
  - 0ℓ (Z→νν): $p_T(Z) > 170\ GeV$
  - 1ℓ (W→ℓν): $p_T(W) > 150\ GeV$
  - 2ℓ (Z→ℓℓ): $p_T(Z) > 150\ GeV$
  - 2ℓ (Z→ℓℓ): $50 < p_T(Z) < 170\ GeV$

- Introduce control regions that closely map each signal region
  - inverted selections → enhance purity in the relevant backgrounds
  - validate analysis variables
  - control / constrain the background normalizations in the fit

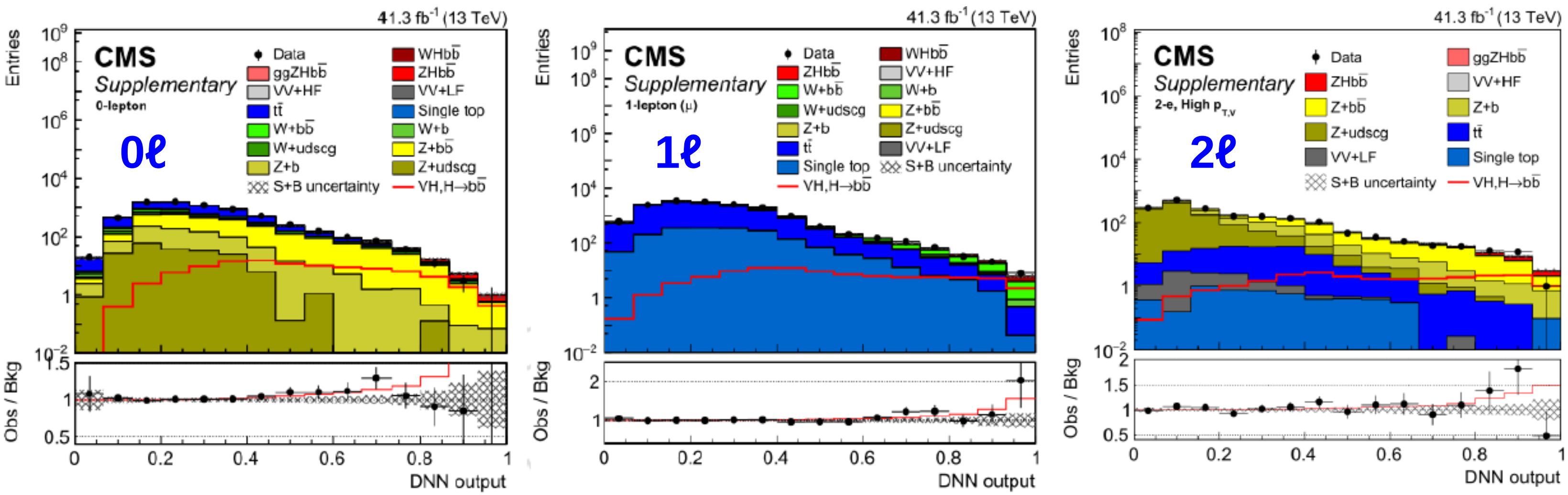
- Simultaneous fit of signal and control regions




[*] Number of additional jets in the event

# Fitting strategy

- Main analysis variable: discriminator of a deep neural network (DNN) to distinguish between signal and background, trained separately in each channel
  - input variables: b-jet properties, di-jet kinematics, event topology
  - carefully validated through data/MC comparison

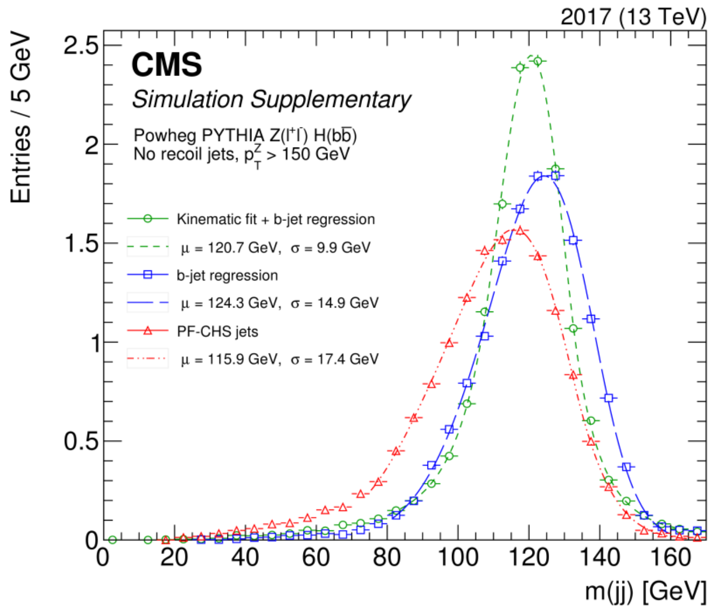



- Additional variables from control regions to exploit different background shapes
  - particularly important for the V+b(b) backgrounds
  - in most cases, use shape of b-tag discriminator or yields
  - 0ℓ and 1ℓ channels: use dedicated DNN multi-categorizer for heavy flavor control region

# Improving mass resolution

- b-jet energy regression:
  - mainly recovers missing energy in jet due to neutrino in semi-hadronic decays
  - recently switched from BDT to DNN algorithm
  - extended set of input variables (including lepton flavor, jet mass, energy fraction in ΔR rings
  - significant mass resolution improvement without background sculpting (→ 11.9% in 2017)

- Kinematic fit:
  - no intrinsic missing energy in Z→ℓℓ channel
  - apply kinematic fit
    - constrain ℓℓ system to Z mass
    - balance ℓℓ + bb system in transverse plane
  - improvement of up to 36% in $m_{bb}$ resolution



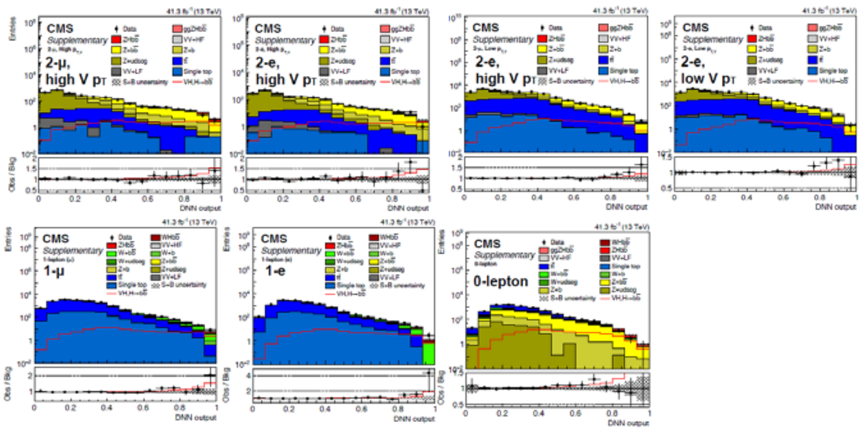
# Systematic uncertainties

- Total uncertainty ~34%
- Major sources of systematic uncertainty:
  - BG normalization & modeling, b-tagging, MC sample size

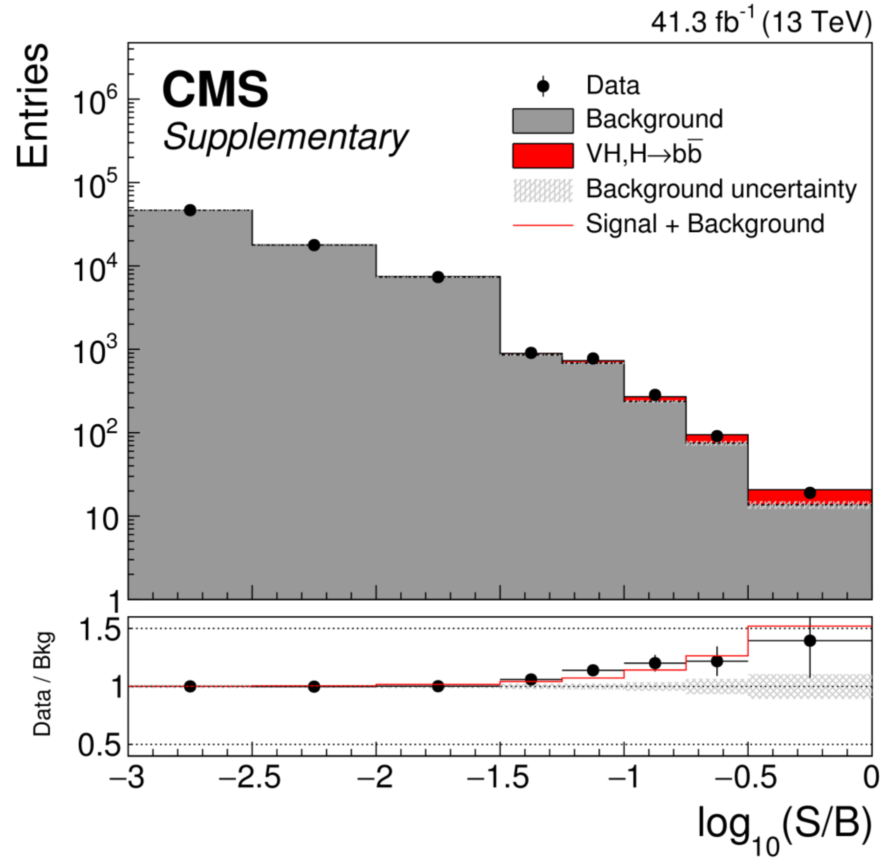
| Uncertainty source | $\Delta\mu$ | |
|---|---|---|
| Statistical | +0.26 | −0.26 |
| Normalization of backgrounds | +0.12 | −0.12 |
| Experimental | +0.16 | −0.15 |
| b-tagging efficiency and misid | +0.09 | −0.08 |
| V+jets modeling | +0.08 | −0.07 |
| Jet energy scale and resolution | +0.05 | −0.05 |
| 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 |

# Signal regions (2017 data)

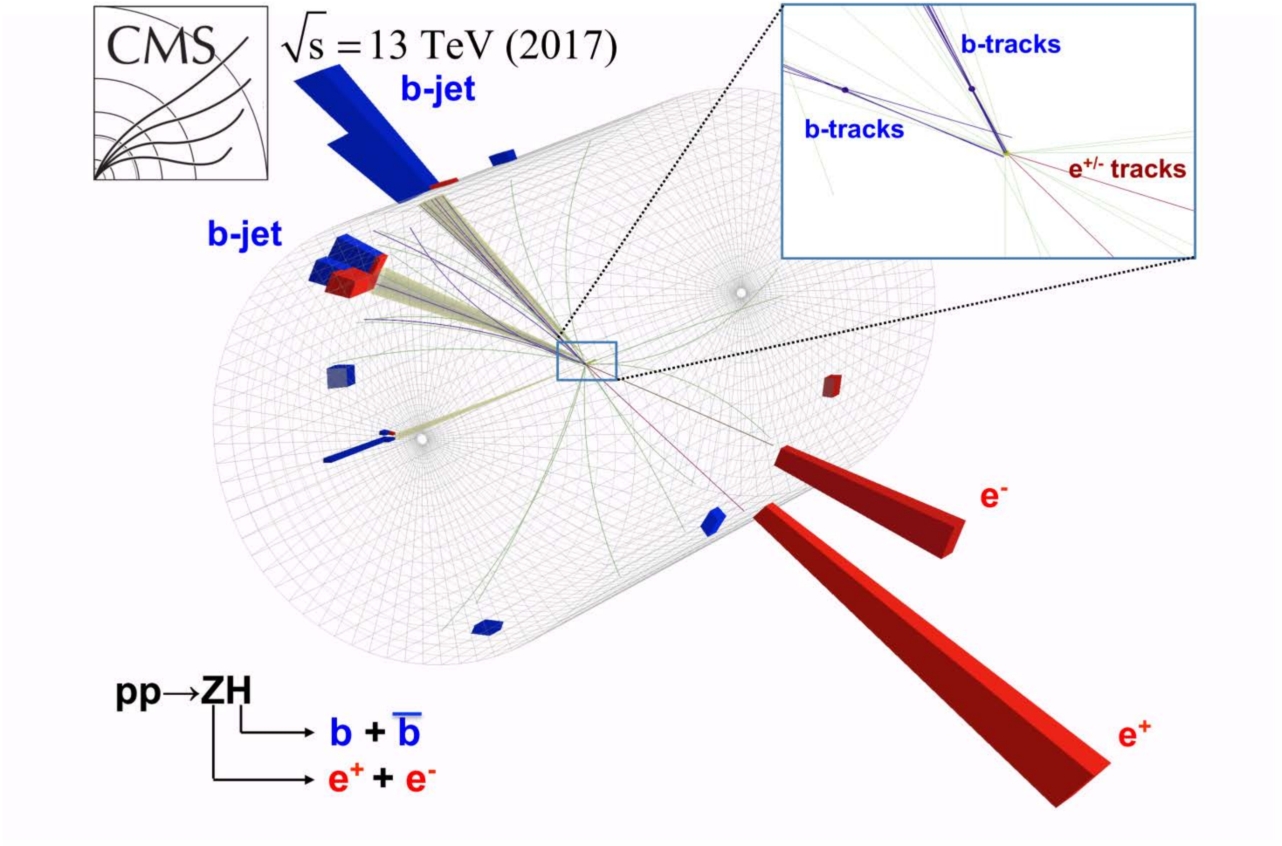
- Sort DNN distributions into bins of similar S/B ratio and combine



- Observe an excess which is clearly compatible with the expected VH(bb) signal



# ZH candidate event display

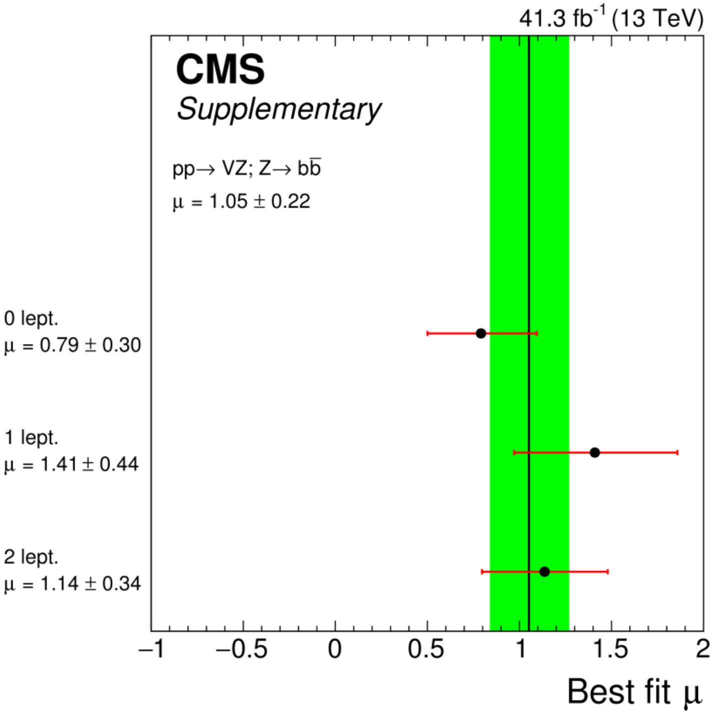


# Validation: VZ(→bb)

- Z production in association with V provides "standard candle"
- Perform VZ analysis with similar technique as VH
  - same DNN inputs but dedicated training
  - larger m(bb) window in signal region → fully include Z(bb) peak

- Results:
  - Signal strength:
    - $\mu = 1.05 \pm 0.22$
  - Significance:
    - 5.2 $\sigma$ observed
    - 5.0 $\sigma$ expected

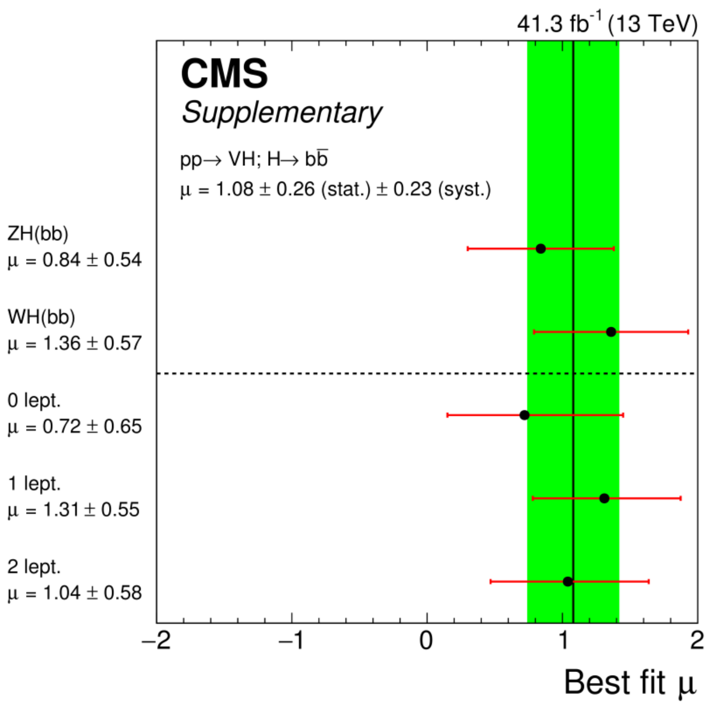
- Confirms proper functioning of methodology



# VH(bb) results with 2017 data

- Very good channel compatibility
- Results from 2017 data agree well with SM expectation
  - signal strength: $\mu = 1.08 \pm 0.34$
  - significance: 3.3 $\sigma$ observed (3.1 $\sigma$ expected)
  - ~5-10% sensitivity increase compared to 2016

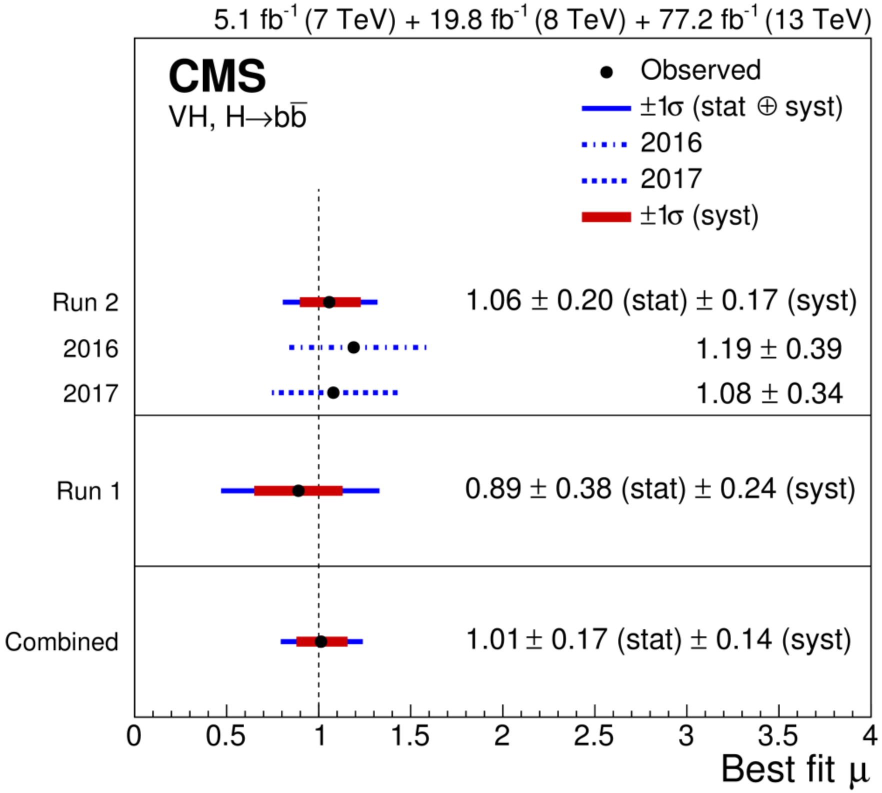
| Data set | Expected | Observed | Signal strength |
|---|---|---|---|
| 2017 | | | |
| 0-lepton | 1.9 | 1.3 | $0.73 \pm 0.65$ |
| 1-lepton | 1.8 | 2.6 | $1.32 \pm 0.55$ |
| 2-lepton | 1.9 | 1.9 | $1.05 \pm 0.59$ |
| Combined | 3.1 | 3.3 | $1.08 \pm 0.34$ |



# Combination of VH(bb) results

- Combine the 2016 + 2017 + the Run 1 analyses

| Data set | Significance ($\sigma$) Expected | Observed |
|---|---|---|
| 2017 | | |
| 0-lepton | 1.9 | 1.3 |
| 1-lepton | 1.8 | 2.6 |
| 2-lepton | 1.9 | 1.9 |
| Combined | 3.1 | 3.3 |
| Run 2 | 4.2 | 4.4 |
| Run 1 + Run 2 | 4.9 | 4.8 |

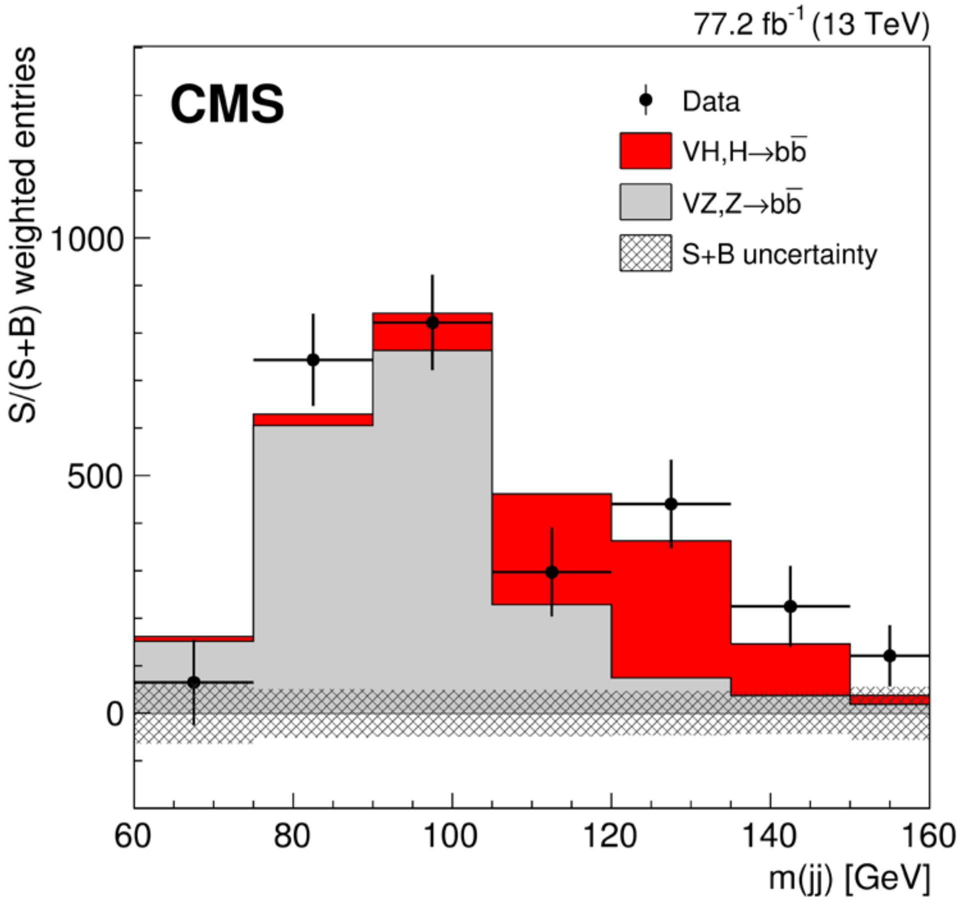



- Close to "observation" level
- Signal strength: $\mu = 1.01 \pm 0.23$

# Visualization of excess: m(bb)

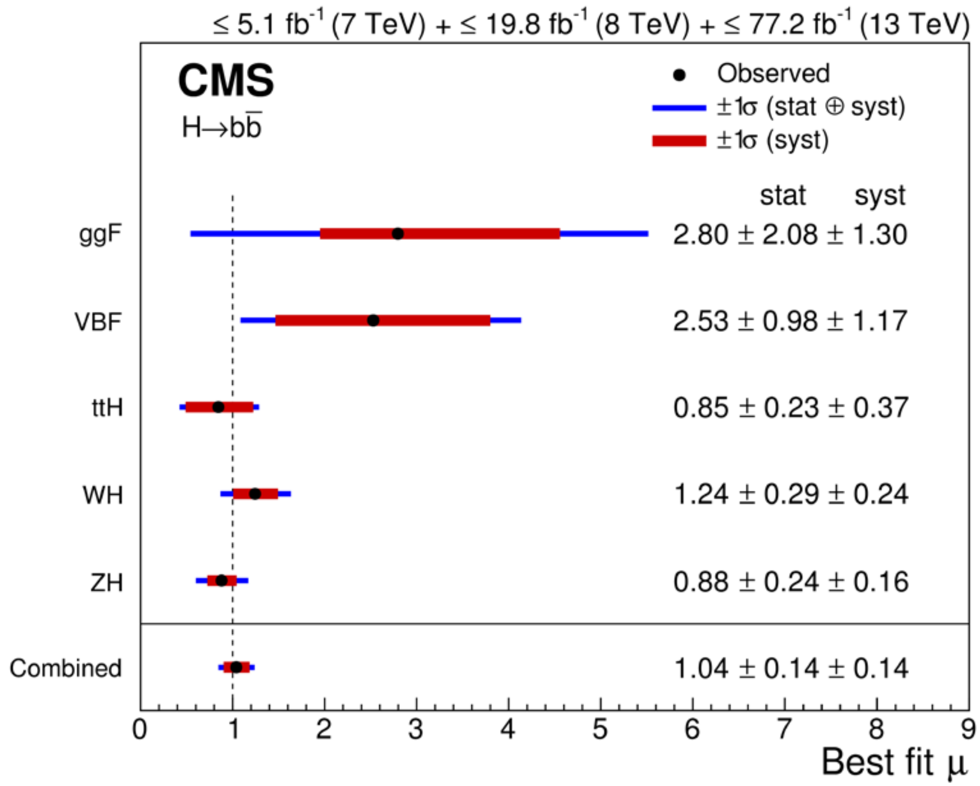
- Categorize events with DNN but remove variables correlated with m(jj)
- Fit to m(jj) distribution
  - lower sensitivity than DNN fit, but direct visualization of signal
- Weight with S/(S+B) and subtract background

- Excess compatible with the sum of VZ(bb) and VH(bb) peaks
  - signal strengths compatible with main analysis



# Combination of H(bb) results

- Combine VH(bb) results with:
  - ttH(bb)
  - boosted ggH(bb)
  - VBF H(bb)
- Measured signal strength:
  - $\mu = 1.04 \pm 0.20$
- Significance:
  - 5.6 $\sigma$ observed
  - 5.5 $\sigma$ expected
- Observation of H→bb decay



# Summary

- CMS achieved 5.6 $\sigma$ observation of H→bb decay
- Signal strength in perfect agreement with SM
- Establishes the dominant decay channel of the Higgs boson
- Confirms the SM picture of Higgs Yukawa couplings at the present level of accuracy
- Completes the set of (currently) accessible couplings to heavy fermions (b, $\tau$, t)

➔ Opens a new era of Higgs precision measurements at the LHC