Higgs decays to third-generation fermions at CMS

Higgs Couplings 2019 - Oxford (UK) - 01.10.2019

Janek Bechtel on behalf of the CMS collaboration

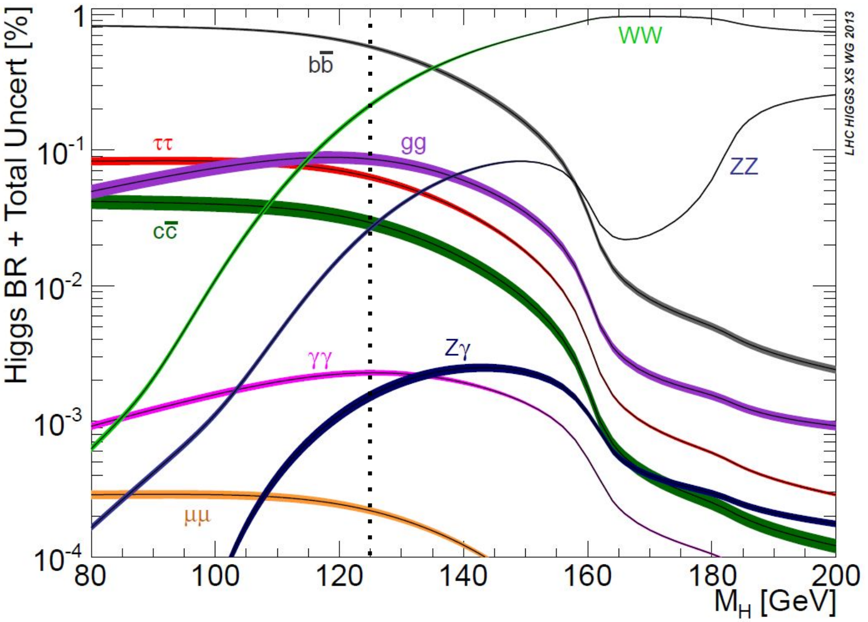
Higgs Decays To Fermions

- Direct probe of the Yukawa coupling - does the Higgs boson couple to fermions as predicted in the Standard Model?

- Branching fraction of the Higgs boson to fermions proportional to particle masses

- Decays of the Higgs bosons into **b-quarks** and **tau leptons** are most sensitive to probe Higgs decays into fermions, however both have their challenges!

- This talk: **SM measurements H→bb** and **H→ ττ**
  - ttH: Today at 14:20



Higgs Decays To Fermions: H→bb and H→ττ

- **We have a clear observation** of the Higgs decay into b quarks (*) and into tau leptons (**) and are now entering the **precision era** at both the **theory** and the **experimental level**

- So far no evidence for BSM physics at ~1 TeV energy scales
  → New phenomena might only appear at larger (out-of-reach) energy scales
  → Use of effective field theories to probe **deviations in the interactions of known SM particles** resulting from **possible BSM particles at an inaccessible energy scale**

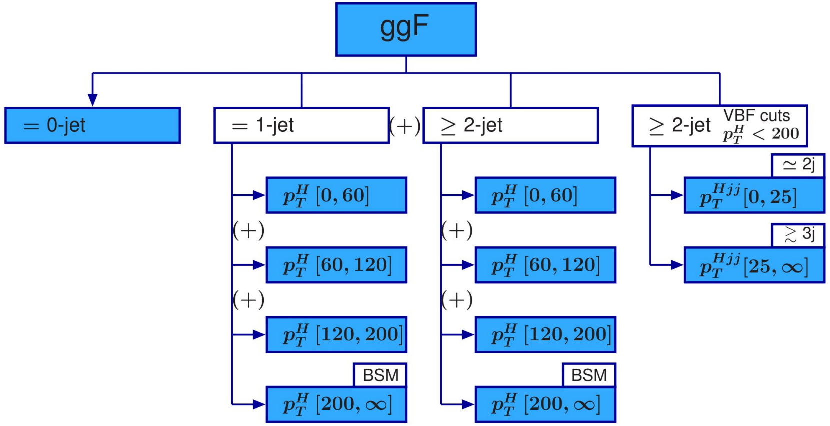
- Precision measurements of the Higgs sector is an obvious tool to search for these deviations

- We are now beginning to **stress-test the Higgs couplings to fermions**

(*) *Phys. Rev. Lett. 121 (2018) 121801*
(**) *Phys. Lett. B 779 (2018) 283–316*

Precision Era Measurements: Simplified Template Cross Sections

- Measured signal strength will no longer be the ultimate figure of merit for analysts
- Simplified template cross sections (STXS) evolve future measurements towards measuring **cross sections in mutually exclusive regions of the phase space** **("STXS bins")**
- Possible BSM effects or deviations in effective field theories can be parameterized into STXS bins
- H→bb and H→ττ provide **access to all Higgs production modes**, with **high sensitivity to VH and VBF**



STXS Stage 1.0 binning for gluon fusion production

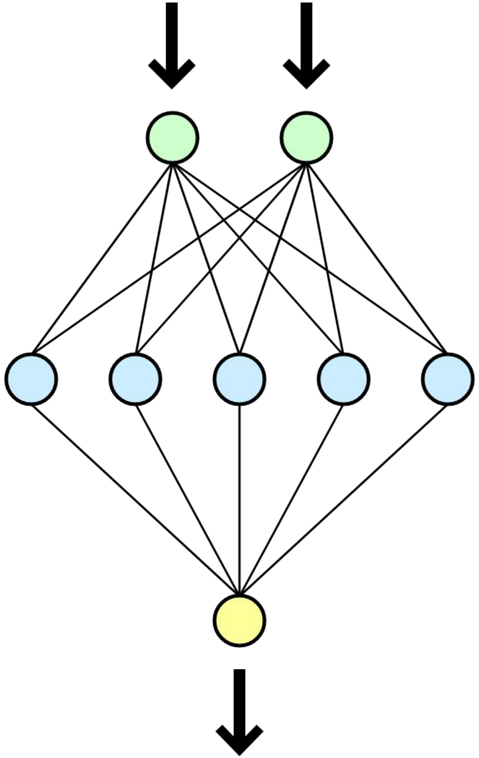
Precision Era Measurements: New Experimental Methods

- MVA techniques are now widely used at various steps of the H→bb and H→ττ analyses, e.g. for
  - identification of b jets or hadronic tau leptons,
  - mass regression or
  - signal vs background discrimination

  → Use of **MVA score as final discriminator** instead of m_{jj} / $m_{\tau\tau}$

  → Discovery of H→bb was possible by the use of modern MVA techniques!

- CMS actively develops data-driven background estimation methods: In H→ττ around 90% of background events are modeled from data
  - **μ→τ embedding** for genuine di-τ events
  - ***F_F*** **method** for estimation of jets misidentified as hadronic taus



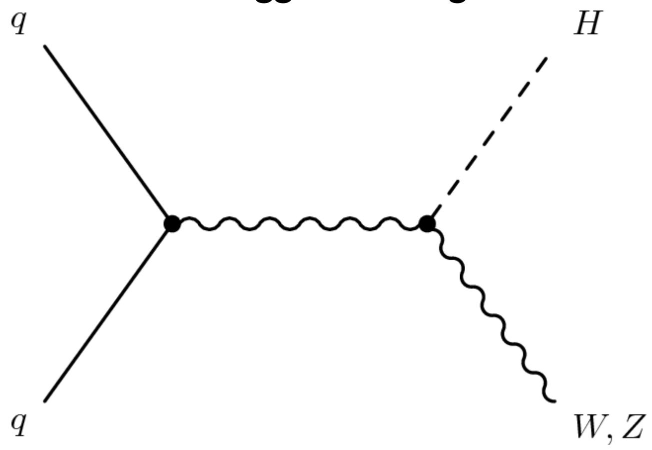
VH(bb)

VH(bb): Overview

most sensitive Higgs production mechanism

- Higgs decay into b quarks is the largest fermionic decay
- **Highest sensitivity of H→bb in VH production**, in which the Higgs boson is produced in association with a W or Z boson
- Leptonically decaying vector boson is helpful for online selection and reduces QCD multijet background
- Most sensitive channel for H→bb, even though VH production cross section is only third-largest of all Higgs production processes

Higgsstrahlung VH



VH(bb): Analysis Strategy

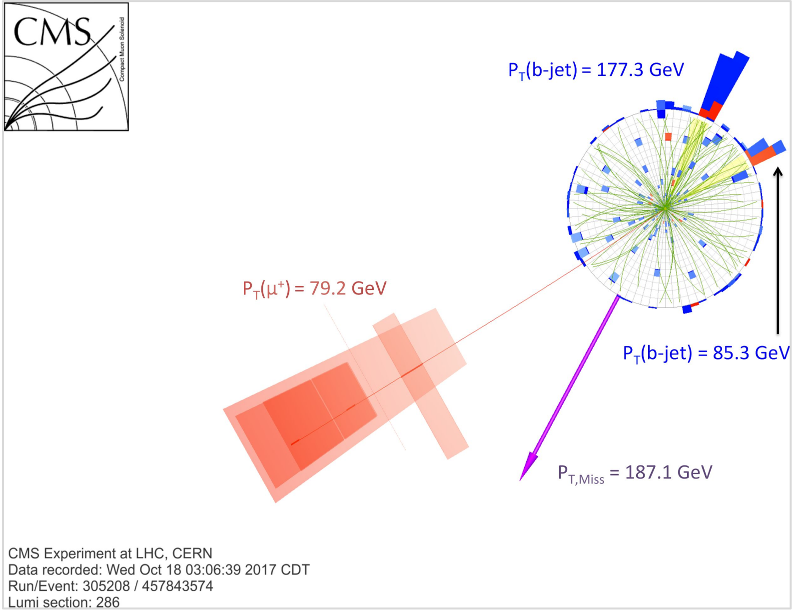
- Selection of events with **0, 1 or 2 electrons or muons** and **2 b-tagged jets**

- Categorization based on 0, 1 or 2 lepton selection and on p_T of W/Z boson

- Most important backgrounds are
  - production of W or Z bosons in association with jets (V+jets),
  - production of top-quark pairs (tt),
  - single top production,
  - diboson (WW, WZ or ZZ) and
  - QCD multijet events

- **Deep neural network** and **kinematic fit in two-lepton channel** improve mass resolution by ~15%

Phys. Rev. Lett. 121 (2018) 121801

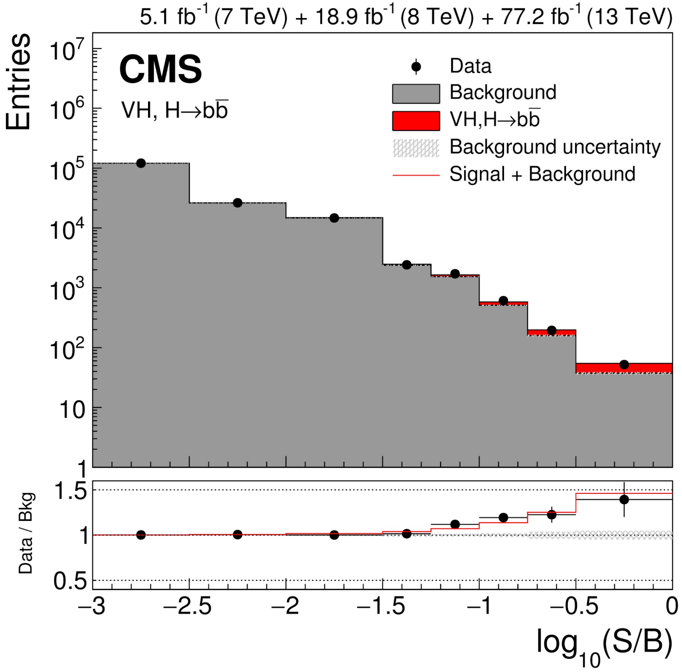


A candidate event for the production of a W boson with a Higgs boson in CMS detector. The Higgs boson decays to two bottom quarks.

VH(bb): Analysis Strategy

- Final discriminator derived by **deep neural network classifier** with 5 hidden layers
- 11-14 input variables depending on W/Z decay channel. Among the most relevant variables are the **b-jet classifier** (deepCSV), **di-jet mass and p_T(V)**
- Separate training of the classifier is performed for each lepton category

Phys. Rev. Lett. 121 (2018) 121801

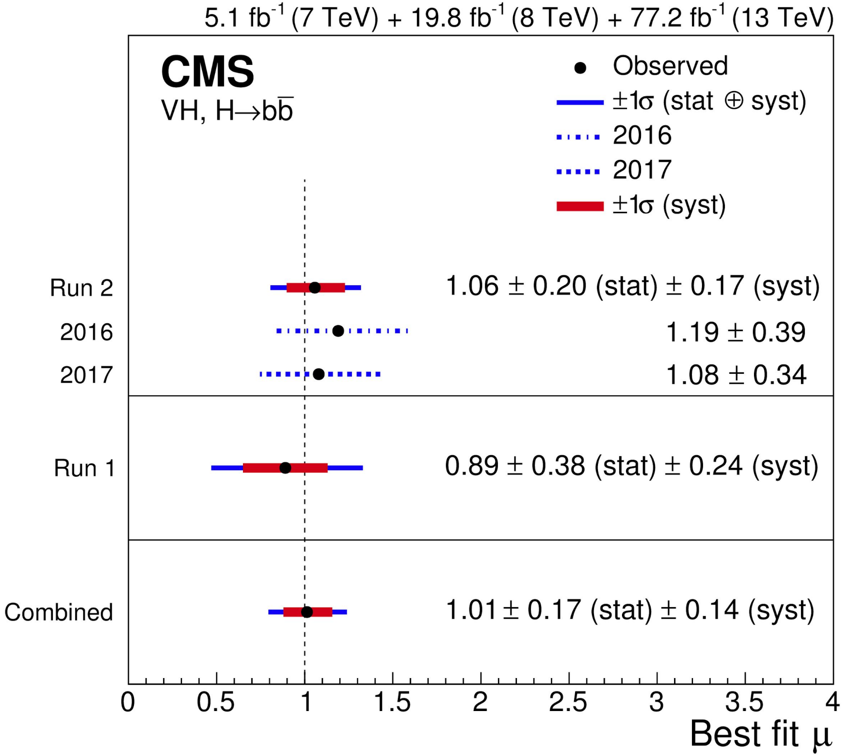


Distributions of event yields sorted into bins of similar signal-to-background ratio, as given by the result of the fit to their corresponding multivariate discriminant.

VH(bb) Results

- The analysis reported a measurement of the signal strength of $\langle\mu_{VH}\rangle = 1.06^{+0.26}_{-0.26}$ for 2016+2017
- In combination with the less sensitive production modes (boosted ggH, ttH, VBF) and results from Run 1, the decay H→bb was observed beyond 5σ last year
- Result for full Run 2 with 137 fb^{-1} is on the way

Phys. Rev. Lett. 121 (2018) 121801



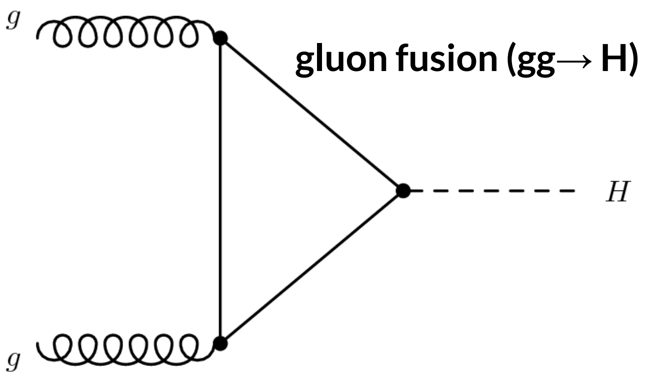
$gg \to H(bb)$

boosted gg→H(bb): Overview

- Both the Higgs production via gluon fusion and the Higgs decay into b quarks have the highest production cross sections and branching fraction respectively

- Still, the gg→H(bb) production is almost impossible to find due to the overwhelming QCD multijet background (at least for unboosted Higgs bosons)

- Analysis focuses on very high Higgs p_T > 450 GeV

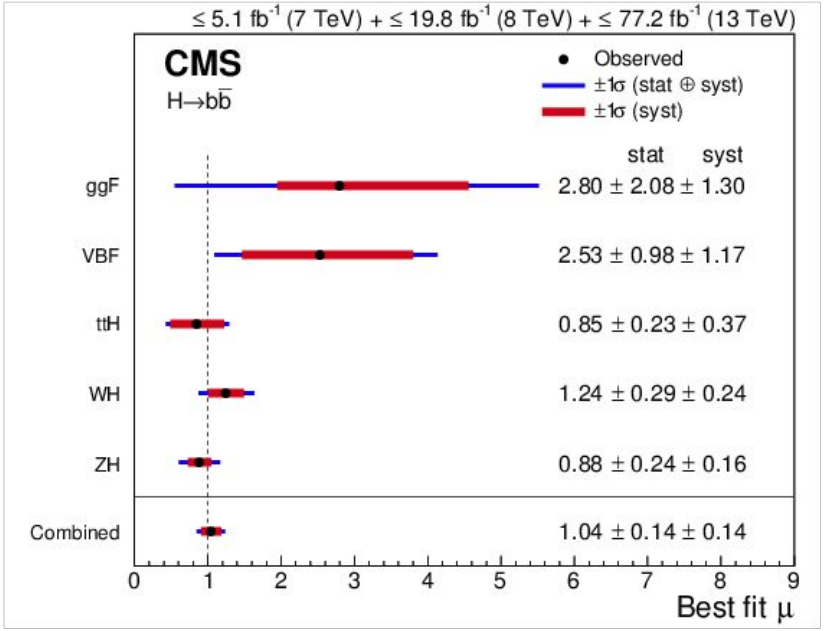


gluon fusion (gg→ H)

boosted gg→H(bb): Results

- Much lower significance than VH(bb) due to high QCD multijet background
- Highly boosted Higgs topology will be useful for **measurements of high-p_T(H) STXS bins**

Phys. Rev. Lett. 121 (2018) 121801



gg→H(ττ) and VBF(ττ)

gg→H/VBF($\tau\tau$)

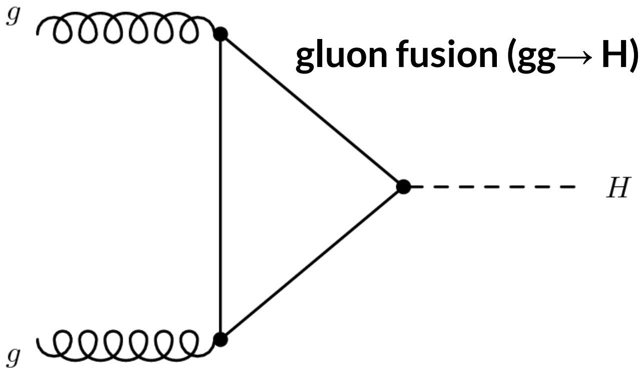
- Higgs decay into tau leptons has the second largest branching fraction of fermionic decays

- **Among all Higgs decays: Highest experimental sensitivity to vector boson fusion production** (VBF)

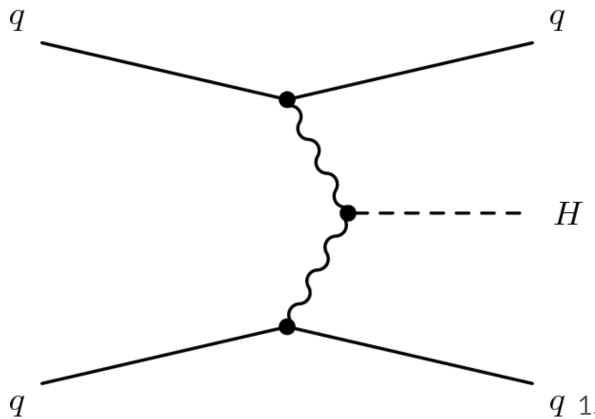
- Latest public result *CMS-PAS-HIG-18-032* using 2016+2017 data relies on several new analysis methods with respect to observation paper, most notably:
  - Large-scale use of **data-driven background estimation methods**
  - The use of a **neural net classifier** for the discrimination of signal and background events
  - Measurement in **bins of STXS**

most sensitive Higgs production mechanisms

gluon fusion (gg→ H)



vector boson fusion (VBF)

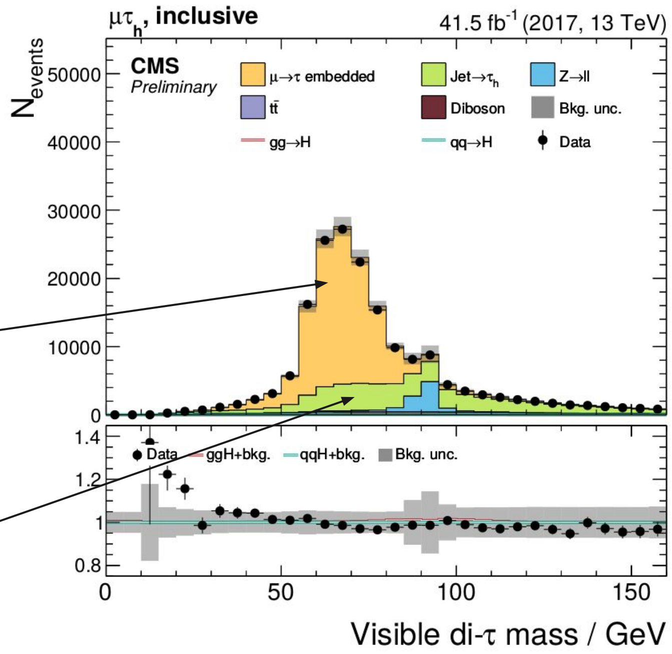


gg→H/VBF(ττ) Background Estimation

new

- Di-τ analysis uses four most sensitive final states: full-hadronic ($\tau_h\tau_h$), semi-leptonic ($e\tau_h$, $\mu\tau_h$), leptonic ($e\mu$)
- **~90% of background events** are estimated from data

Process	**Estimated by ...**
genuine ττ events (Z→ττ, tt→ττ, VV→ττ, ...)	**μ→τ embedded events** **(data-driven)**
prompt leptons + genuine taus (tt→lτ, VV→lτ, ...)	Monte Carlo simulation
jets misidentified as τ_h (QCD, W+jets, tt+jets, ...)	**F_F method** **(data-driven)**
e/μ misidentified as τ_h (Z→ee, Z→μμ, ...)	Monte Carlo simulation



All genuine di-τ (yellow) and jets misidentified as τ_h (green) processes estimated from data

Background Estimation - Embedded Events

new

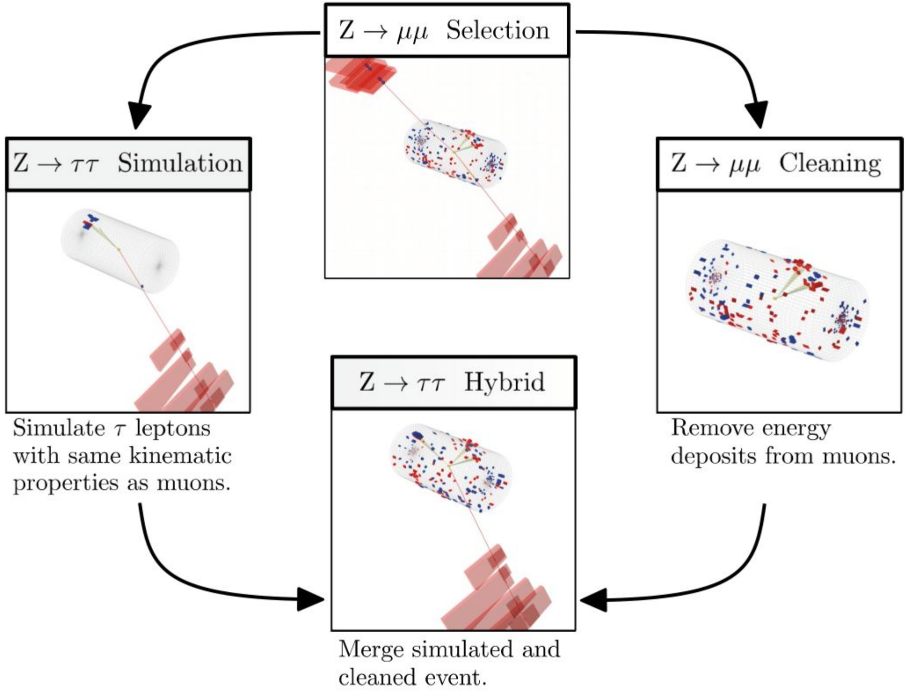
[JINST 14 (2019) P06032](#)

- Z→μμ and Z→ττ decays have the same rate and characteristics

- This enables the **embedding technique,** which allows describing tau backgrounds almost completely from data

- The two **muons are removed** from the reconstructed event record and **replaced by two simulated tau lepton decays**

- **Only the two tau decays are simulated -** No simulation and tuning of pileup required



Background Estimation - F_F method

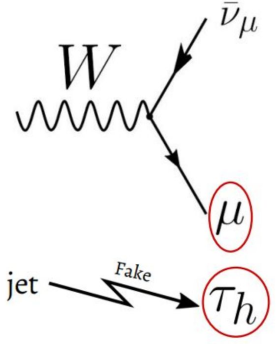
new

JHEP 09 (2018) 007

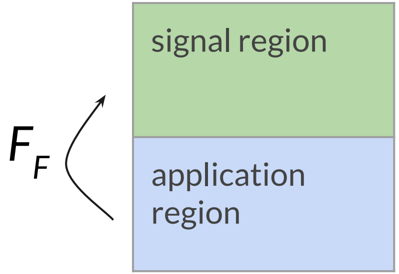
- **Data-driven method of estimating jets misidentified as a τ_h** (in $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$) from
  - W+Jets
  - QCD multijet
  - tt + jets

- 65% of tau leptons decay into hadrons → Challenging to discriminate from QCD background, even with advanced identification methods

- Extrapolation factors for regions given by two orthogonal tau identification requirements are determined for each background process



Tau ID working point:

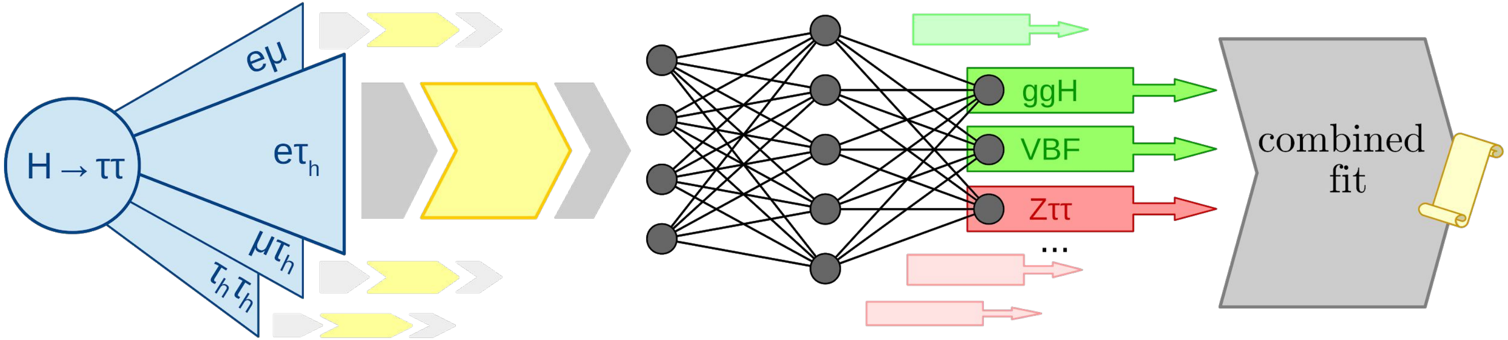


Tight

!Tight && Very Loose

gg→H/VBF(ττ): Analysis Strategy

new



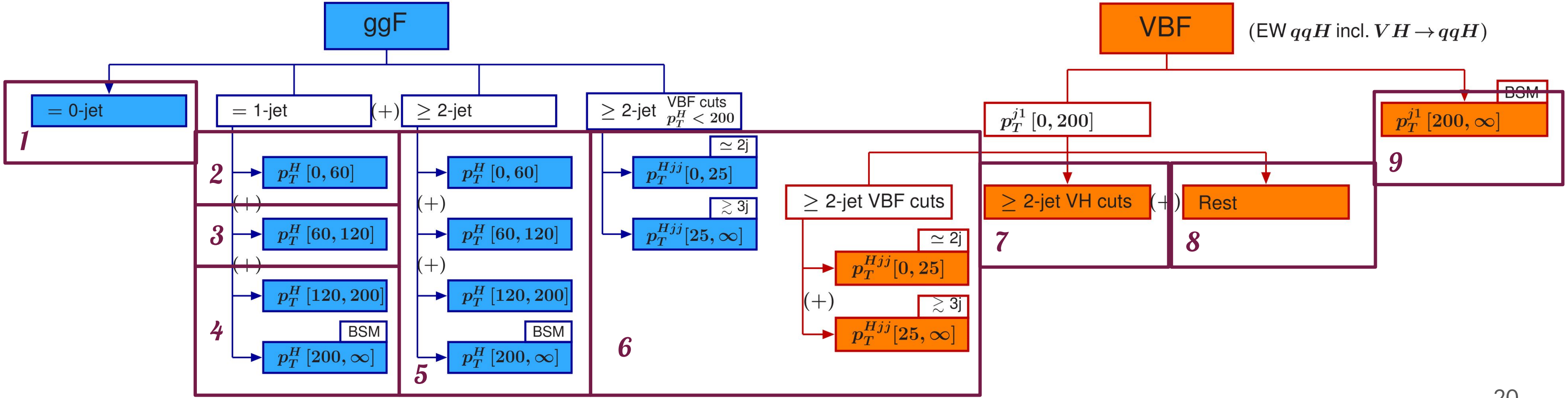
- Neural network is used to classify events into **several signal and background** categories
- Each event will receive a score for each signal and background category. The highest score determines the category of the event
- **Categorization performed not by analyst but by multiclass neural network** - background processes are sorted into background control categories, and signal events into signal categories
- 16-21 event variables depending on decay channel are used as input. Most relevant variables are **di-tau mass, visible di-tau mass, di-jet mass and lepton and jet p_T**

CMS-PAS-HIG-18-032

gg→H/VBF(ττ) Simplified Template Cross Sections

new

- HIG-18-032 results interpreted in **merged Stage 1.0 STXS bins** in 9 separate parameters-of-interest
  (5 for ggH, 4 for VBF as indicated by the boxes below)



gg→H/VBF(ττ) Results

new

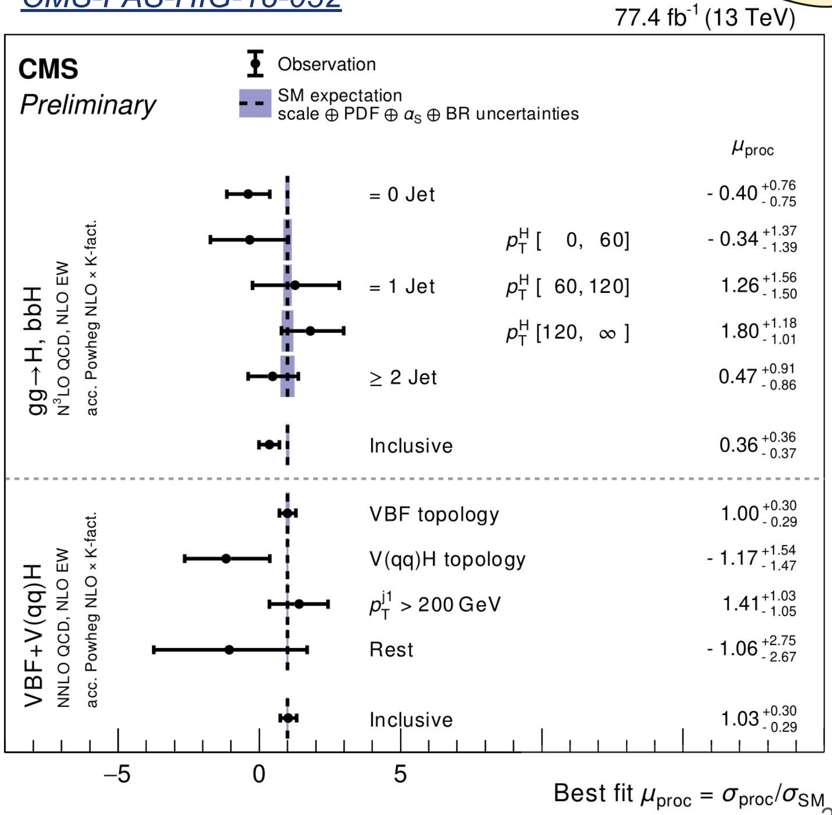
CMS-PAS-HIG-18-032

- The analysis reported the measurement of the inclusive H→ττ signal strength of
  $\langle\mu\rangle = 0.75\,^{+0.18}_{-0.17}$

  and the measurement of the gg→H(ττ) and VBF(ττ) production of
  $\langle\mu_{gg\to H}\rangle = 0.36\,^{+0.36}_{-0.37}$

  $\langle\mu_{VBF}\rangle = 1.03\,^{+0.30}_{-0.29}$

- Result for full Run 2 with 137 fb^{-1} is on the way



VH(ττ)

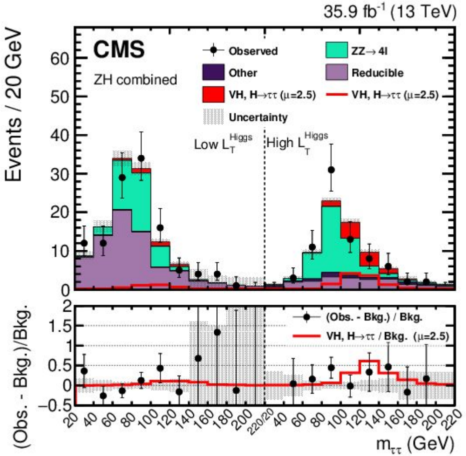
VH(ττ): Analysis Strategy

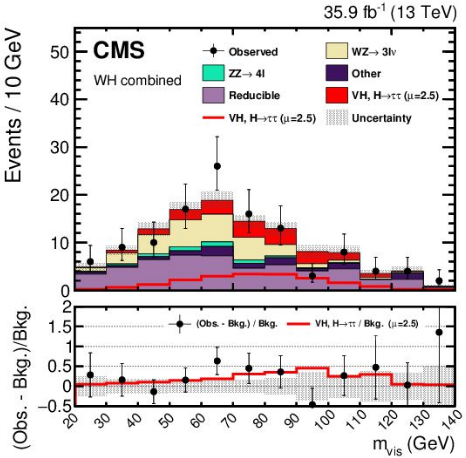
JHEP06(2019)093

- Lower event yield, however also greatly suppressed background due to additional W or Z boson decaying into leptons

- Events with additional ee or μμ from a Z, or additional electron and muon from a W decay are selected

- The latest publication using 2016 data reported a VH(ττ) signal strength of $<\mu_{VH}> = 2.5^{+1.4}_{-1.3}$ with the full Run 2 result on the way





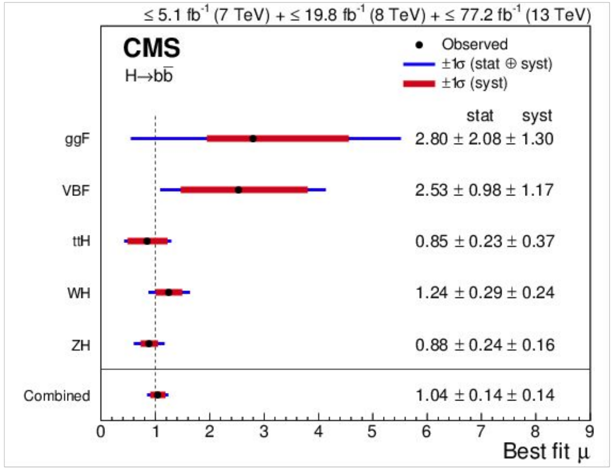
Conclusion

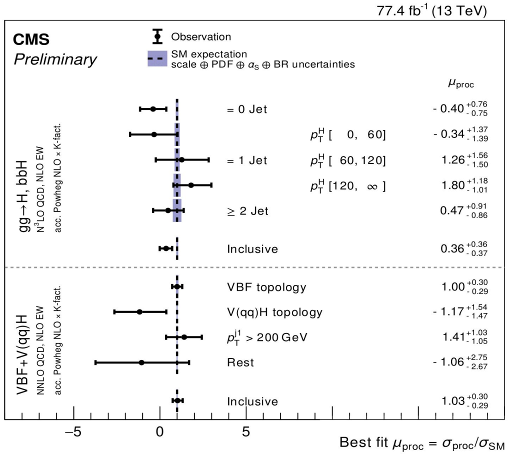
Conclusion

- We are entering the era of **precision measurements** in the analysis of Higgs decays into third-generation fermions

- **Modern MVA techniques such as neural networks** are now widely used in analysis of both H→bb and H→ττ decays

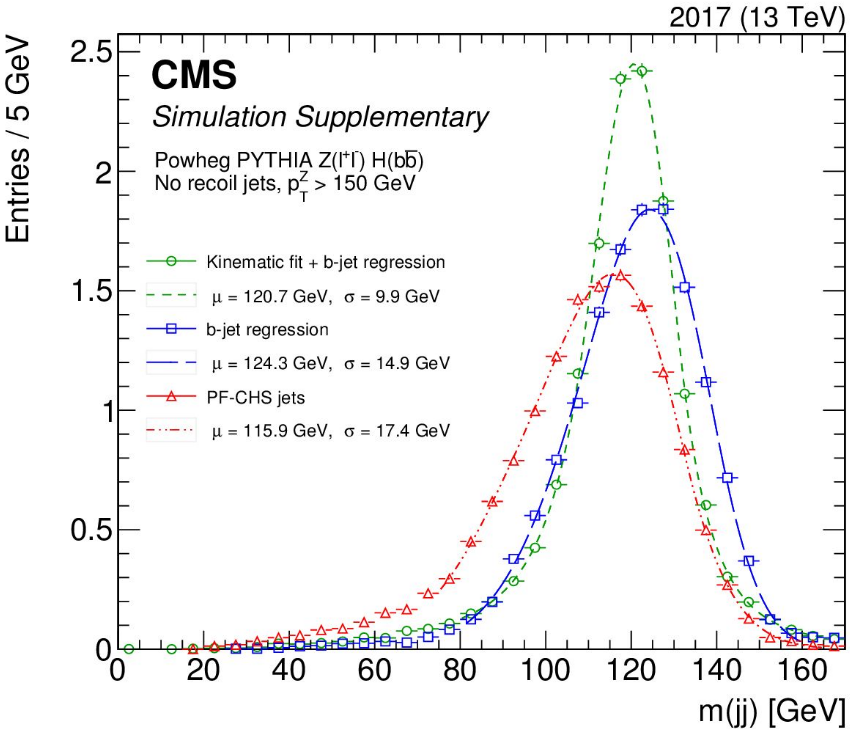
- Total significance no longer the only parameter of interest - Results from the full Run 2 data collected at the CMS experiments will be given in form of **simplified template cross sections**

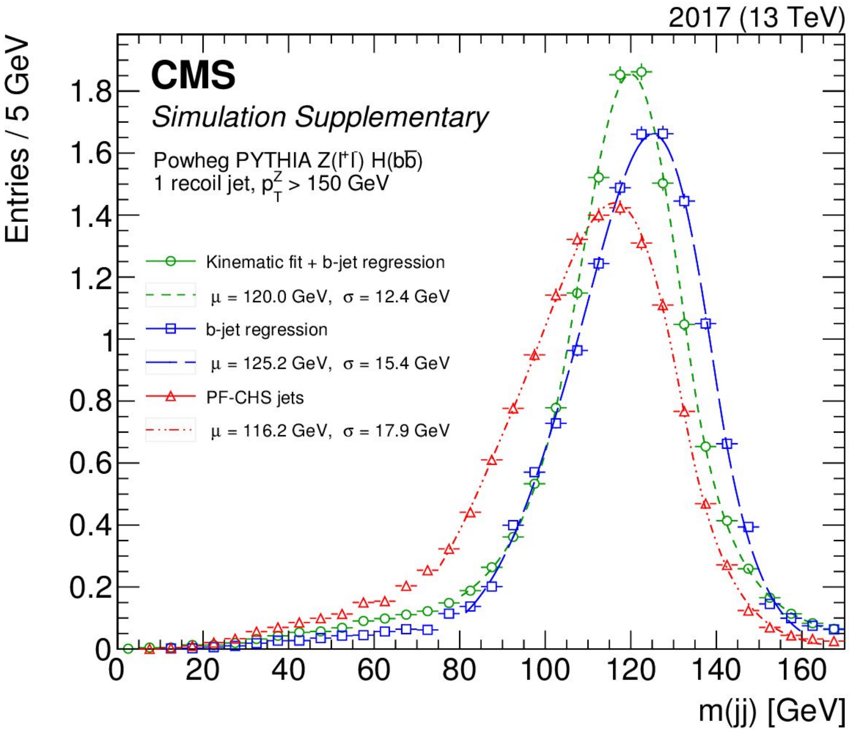




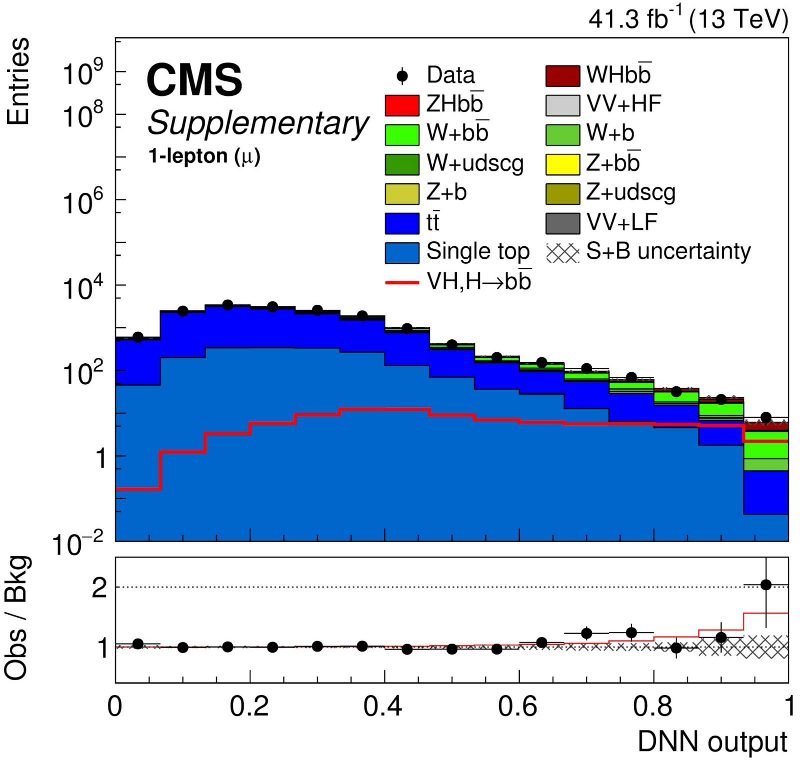
Backup

VH(bb): Mass Resolution

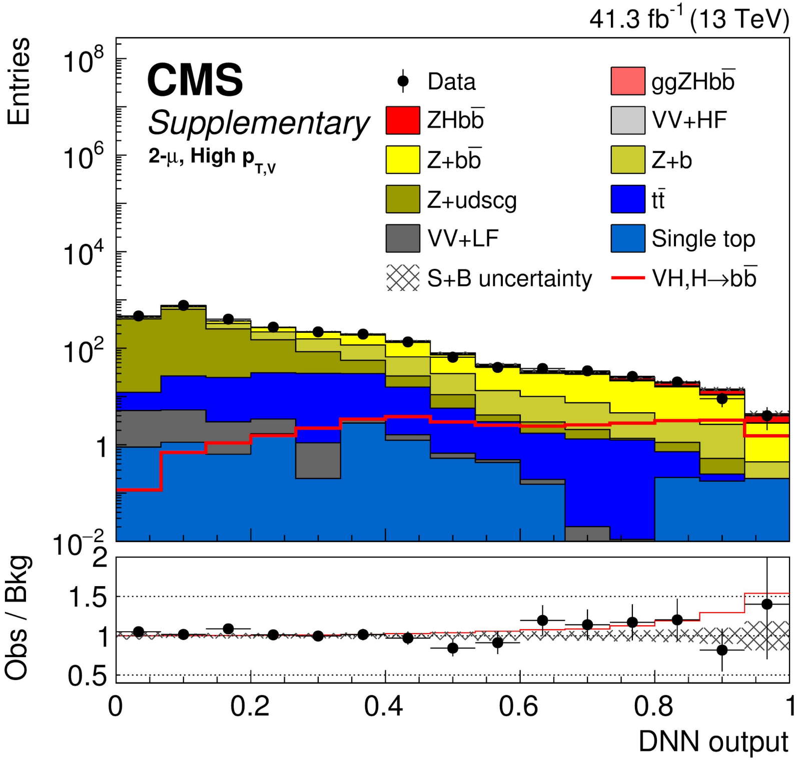




VH(bb): DNN output examples



1-muon category



High pT(V) - 2-muon category

VH(bb): Dominant Uncertainties

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

H→ττ: Background Estimation - F_F method

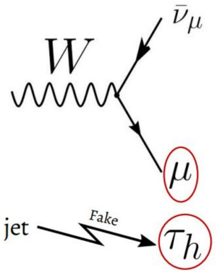
new

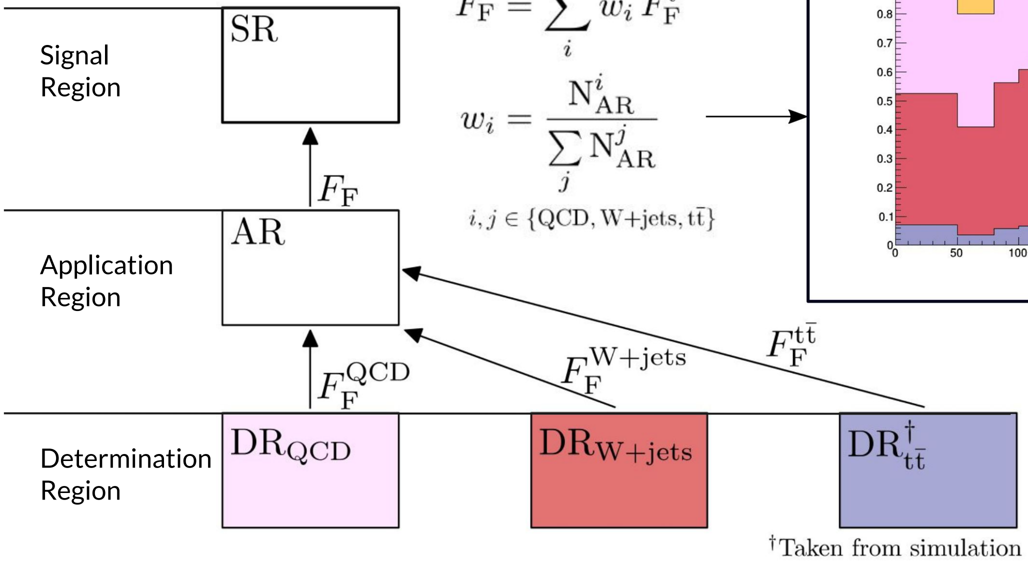
- Data-driven method of estimating jets misidentified as a τ_h (in $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$)
- Used to estimate
  - W+Jets
  - QCD multijet
  - tt + jets

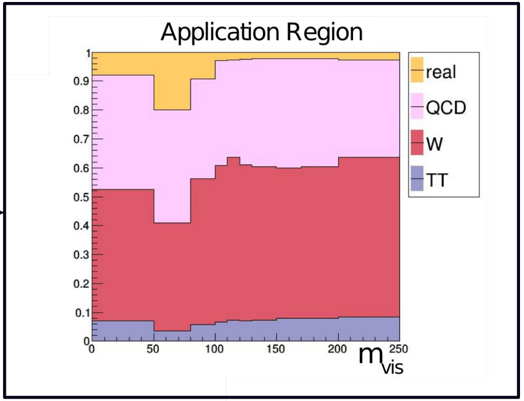
$$F_{\mathrm{F}} = \sum_i w_i \, F_{\mathrm{F}}^i$$

$$w_i = \frac{N_{\mathrm{AR}}^i}{\sum_j N_{\mathrm{AR}}^j}$$

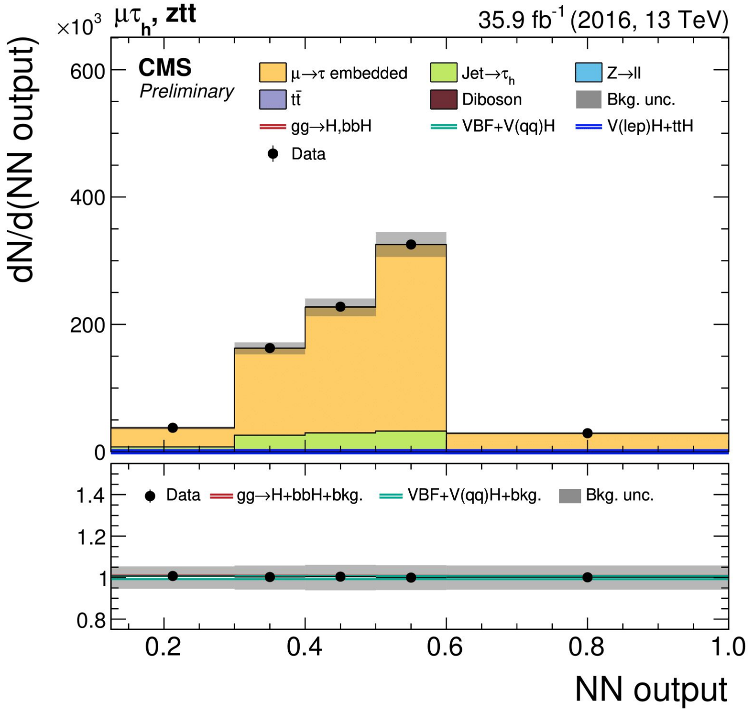
$$i, j \in \{\mathrm{QCD}, \mathrm{W{+}jets}, t\bar{t}\}$$



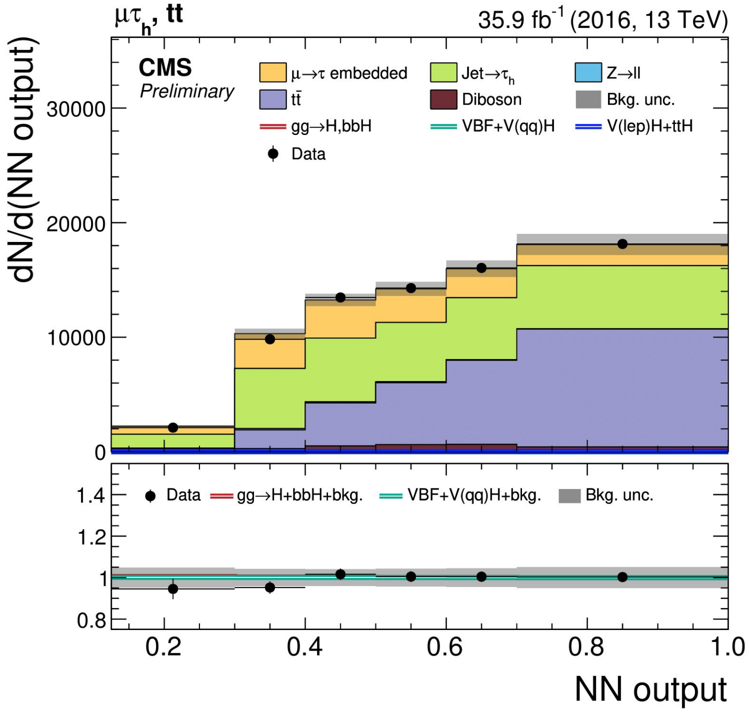




H→ττ: DNN output examples

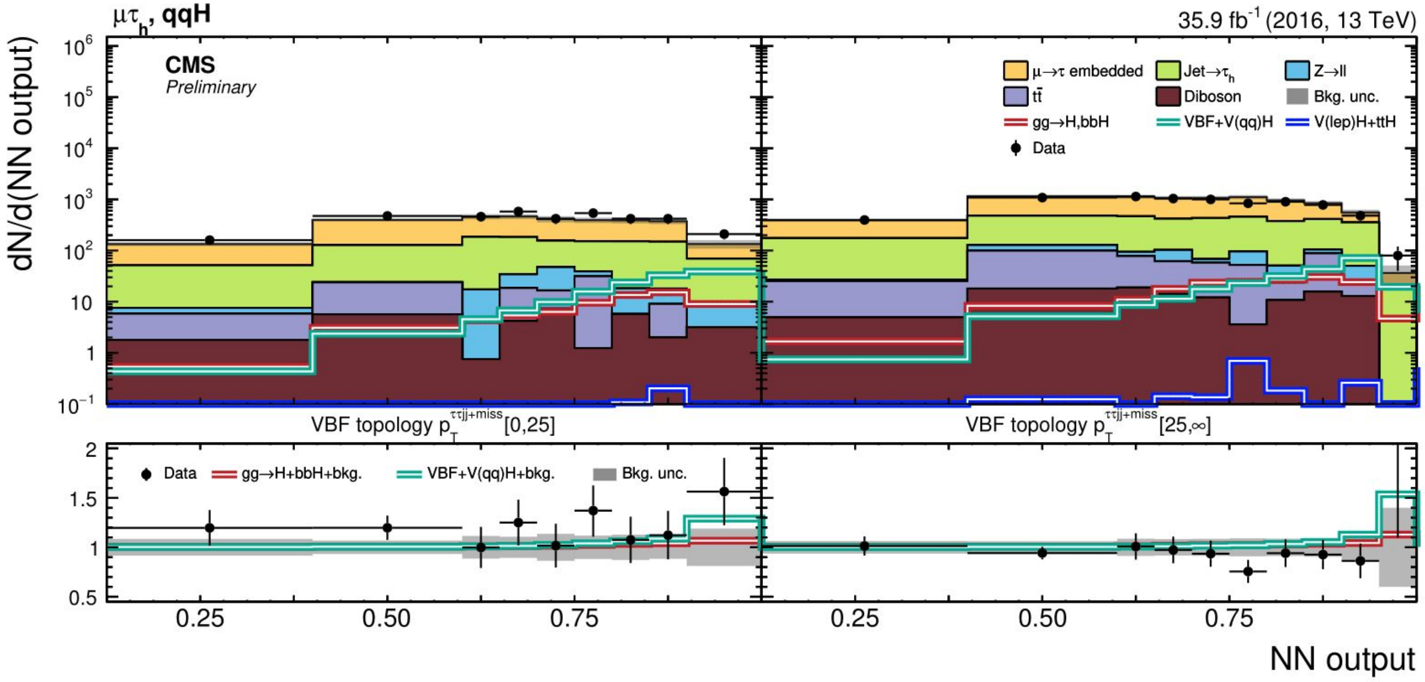


$\mu\tau_h$ channel: Z→ττ background category



$\mu\tau_h$ channel: tt background category

H→ττ: DNN output examples



$\mu\tau_h$ channel: VBF signal category - VBF Topology STXS bins

H→ττ: S/(S+B) binned