Higgs boson measurements in VBF production mode at CMS

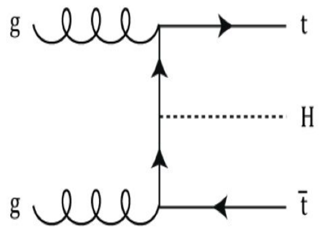
Soumya Mukherjee
TIFR, Mumbai (India)
On behalf of the CMS collaboration

8th Nov, 2022

Sponsored by Infosys-TIFR "Leading Edge" Travel Grant

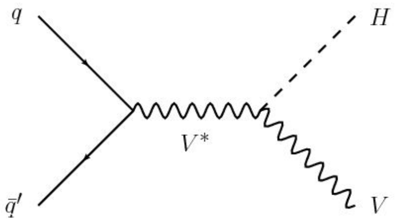
Higgs production mechanisms at the LHC

$t\bar{t}H$



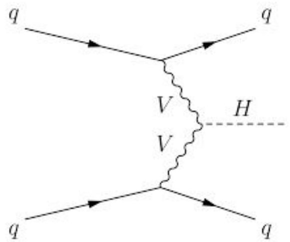
Xsec: 0.5 pb

VH



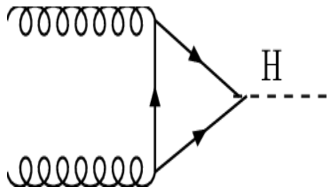
Xsec: 2.61 pb

VBF



Xsec: 3.78 pb

ggF



Xsec: 48.6 pb

Production cross section

Key features of VBF process:

Vector Boson Fusion (VBF) → sub-leading Higgs production process

→ Cross section 3.78 pb @ N^{2}LO QCD and NLO EWK accuracy

→ Very distinctive topology by a pair of forward-backward jets referred to as “VBF” jets with high invariant mass (m_{jj}) and high pseudo-rapidity gap ($\Delta\eta_{jj}$)

Results from Run-2 combination

[Nature 607 (2022) 60](https://www.nature.com/)

→ Combined signal strength of VBF-H process (μ_{VBF}) :

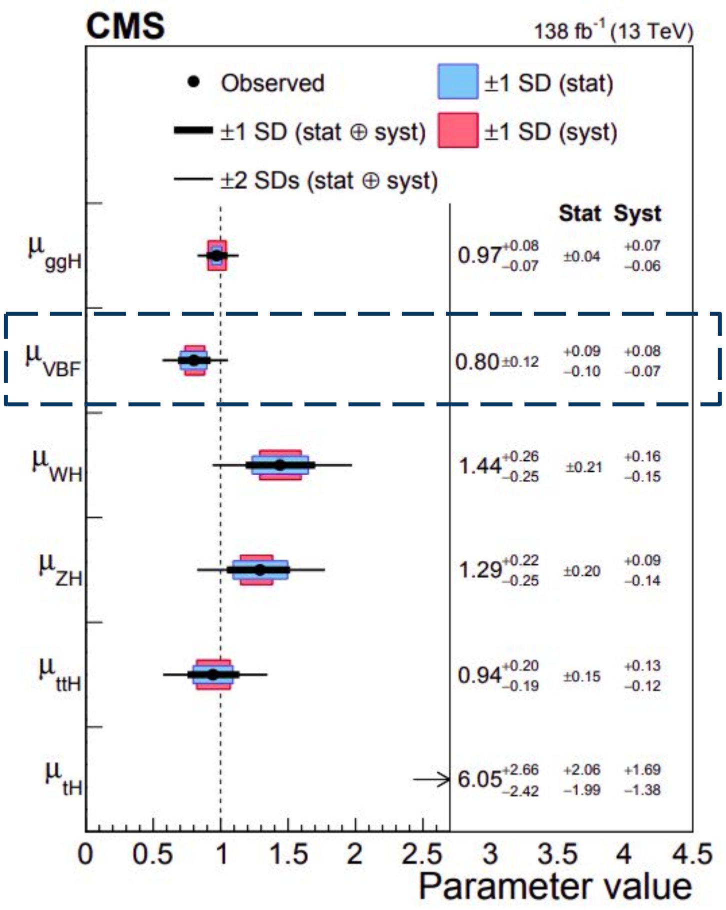
0.80 ± 0.12

→ Observed value is **< 1σ** away from the SM prediction

→ Total uncertainty is equally divided into systematics and statistics

→ All Higgs decays modes are included here **except H → bb**

→ VBFH(bb) in resolved topology is being presented for the first time here from CMS using Run-2 data



Resolved VBF H → bb

CMS PAS_HIG_22_009

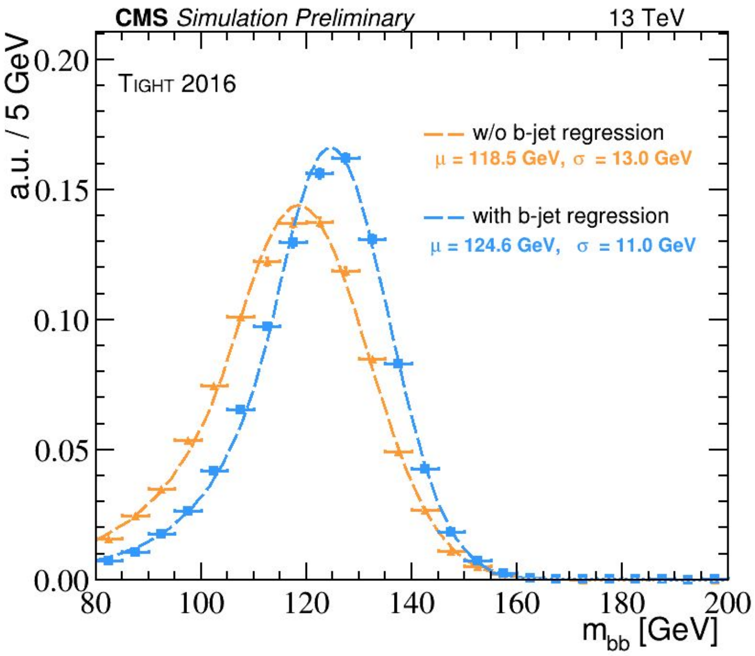
- **Experimentally challenging** due to
  → The overwhelming QCD background events
  → Large resonant Z→ bb background (signal in the higher tail of the Z peak)

- Two dedicated high level trigger [HLT] using **VBF topology and b-tagging requirements**
  TIGHT : stringent VBF criteria with loose b-tag
  LOOSE : Loose VBF requirement with tight b-tag
  → The analysis is performed with the accumulated Run2 data of **91 fb^{-1} integrated luminosity.**

- Two highest b-tagged AK4 jets are selected as Higgs candidate

- **DNN based b-jet energy regression** [Computing and Software for Big Science 4 (2020) 10] applied to improve the mass & resolution of the m_{bb} spectrum
  [15% improvement on σ/μ of signal process]

- Multivariate techniques used to separate signal from the contributing backgrounds



Finally m_{bb} distribution has been used to extract the results

VBF H → bb : Separation of signal from background

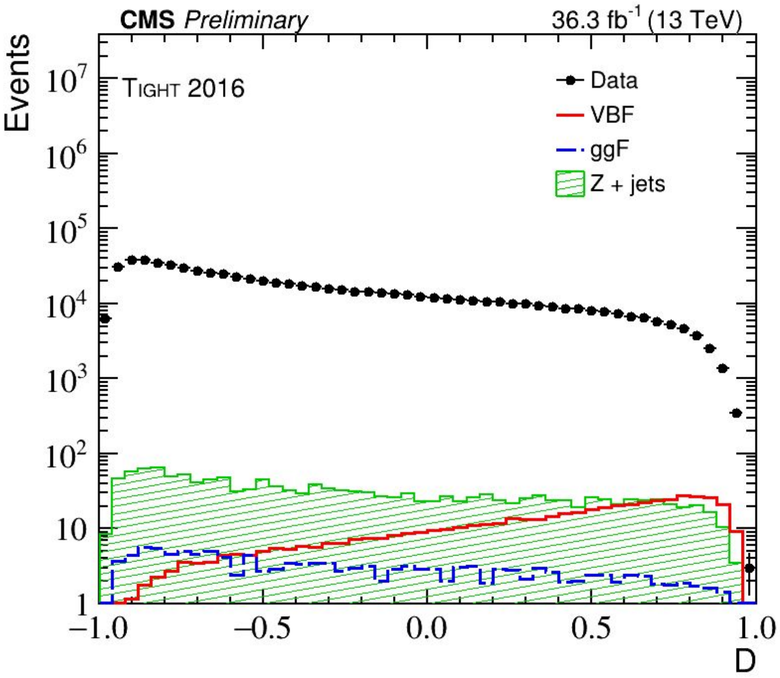
CMS PAS_HIG_22_009

- ❑ Separate BDT trainings
  TIGHT : Binary classifier (VBF vs QCD)
  LOOSE : Multi-classifier (VBF vs ggF vs ZJets Vs QCD)

- ❑ Events are categorized based on discriminants' score (D)

- Categories target not only VBF, but ggF and Z + jets as well
- **ggF categories** improve sensitivity to the inclusive signal strength (ggF + VBF)
- **Zbb categories** establish Zbb standard candle and constrain rate of the Z+jets production

- ❑ Total 18 categories [9 each year]
  → All categories have been used simultaneously to extract the final results



VBF H → bb : Signal and background fit

→ **Signal and resonant background (Z + jets)** estimated from MC

→ Fitting a Crystal Ball function at core

→ 2nd order bernstein polynomial at high tail

→ Shape and normalization of the continuum backgrounds are derived from fit-to-data using Generic polynomial functions (F_i)

$$F_i^{QCD} = \exp\left(-b_i \cdot m_{bb}\right) \cdot \left(1 + \sum_{j=1}^{n} a_{ij} \cdot m_{bb}^{j}\right)$$

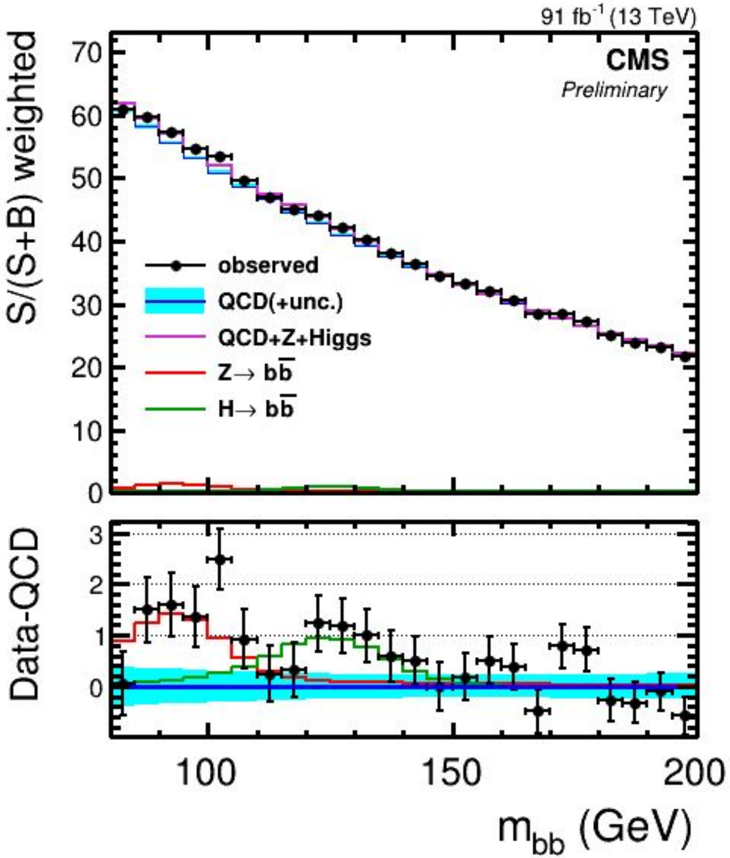
i : number of category, j : order of the polynomial function

→ Order of the polynomial is determined by F-Test

→ Bias study has been performed in each analysis categories by using alternative parametric function to fit data,

→ insignificant bias has been found (vanishing impact on results)

CMS PAS_HIG_22_009



VBF H→bb : Results

CMS PAS_HIG_22_009

signal significance in std. dev.

Inclusive Hbb → VBF + ggF contributions

Process	Observed	Expected
Inclusive Hbb	2.5	2.9
VBFH-bb	2.4	2.7

signal strengths

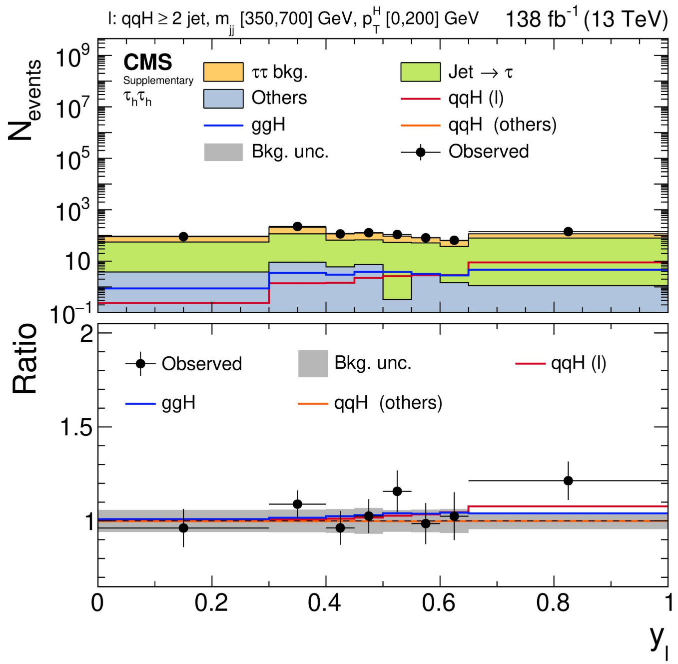
Process	Observed signal strength
μ_{Hbb}	$0.92^{+0.32}_{-0.32}(\text{stat.})^{+0.31}_{-0.22}(\text{syst.})$
$\mu_{VBF\text{-}Hbb}$	$0.97^{+0.35}_{-0.35}(\text{stat.})^{+0.39}_{-0.28}(\text{syst.})$
μ_{Zbb}	0.94±0.20(stat.)±0.21 (syst.)

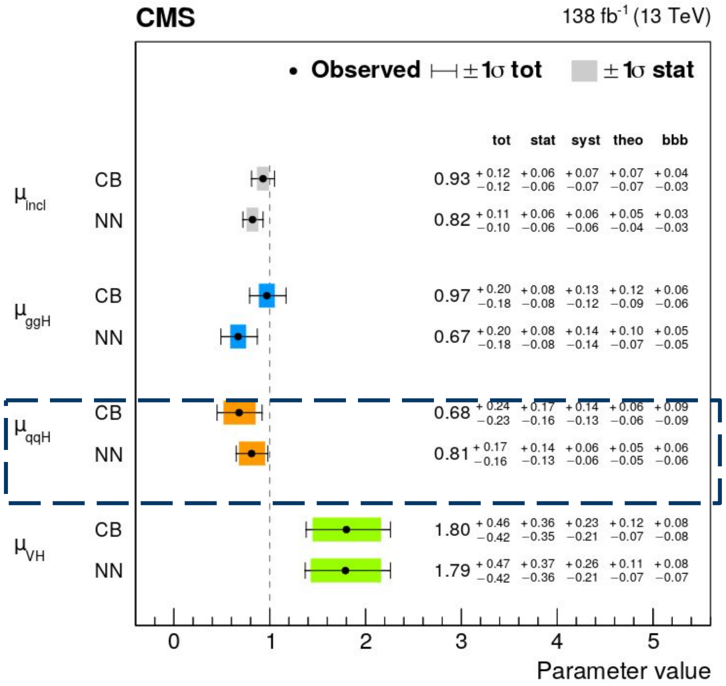
→ Measurement of Zbb performed as a standard candle to validate our analysis
→ Measurements of Hbb, Zbb & VBFH-bb found to be close to the SM prediction [< 1σ]

VBF H→ $\tau\tau$

arXiv: 2204.12957
Submitted to EPJC

→ H → $\tau\tau$ deacy allows direct coupling of Higgs to fundamental fermions
→ $\tau_h\tau_h$, $\tau_h e$, $\tau_h\mu$, $e\mu$ final states considered for ggH & VBF modes
→ **Neural network based DEEPTAU** algorithm used to identify τ_h
→ Analysis is performed in two consecutive way cut based (CB) and Neural network based (NN)

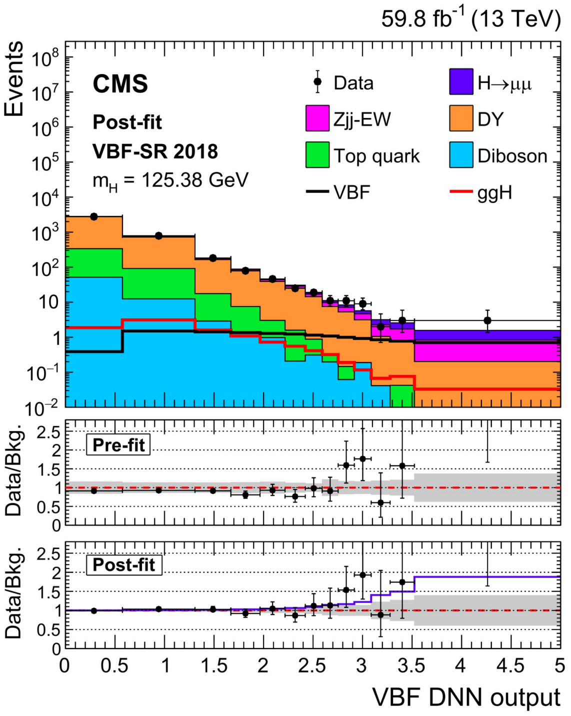


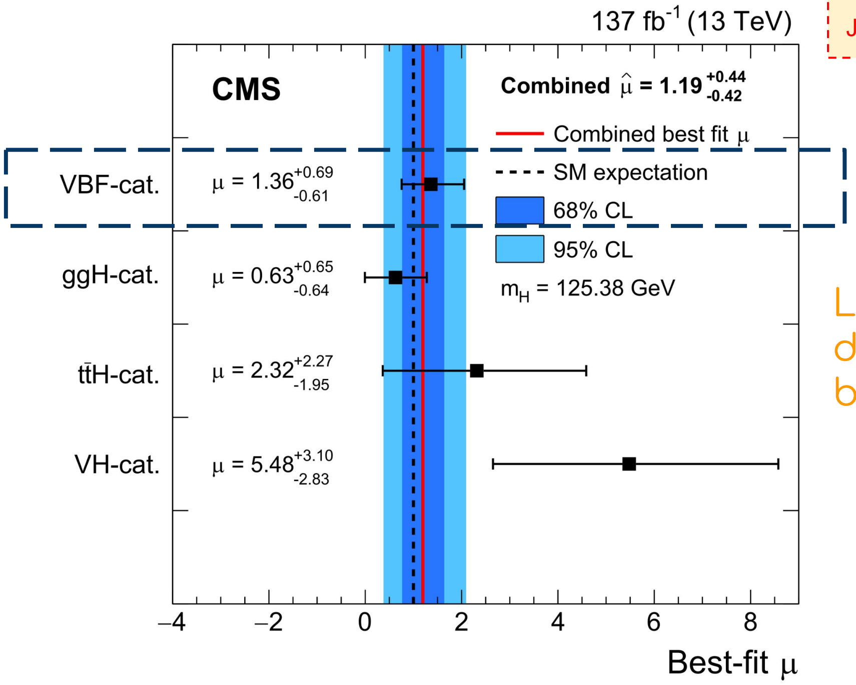


VBF H→ $\mu\mu$

Higgs rare from CMS & ATLAS
→ Giulio Umoret

→ QCD induced DY and EWK Zjj (VBF Z→ $\mu\mu$) are the dominant uncertainties
→ Dedicated **DNN training** (including dimuon mass $m_{\mu\mu}$) has been performed to separate VBF signal from background
→ Results are extracted by fitting **DNN output distribution** for both signal region(SR) and side-band regions (SB)
→ VBF signal sensitivity improved by **20%** from the traditional data-driven background estimation by fitting $m_{\mu\mu}$





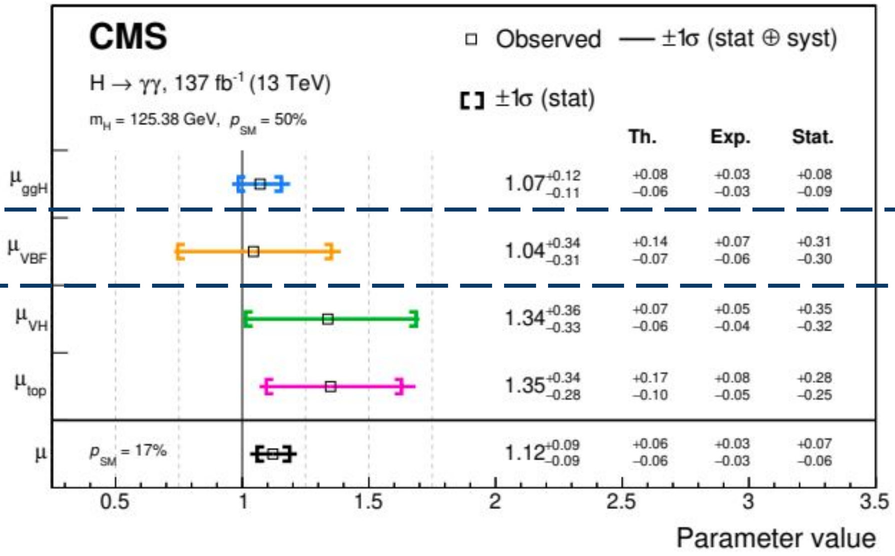
JHEP 01 (2021) 148

Largely dominated by statistics

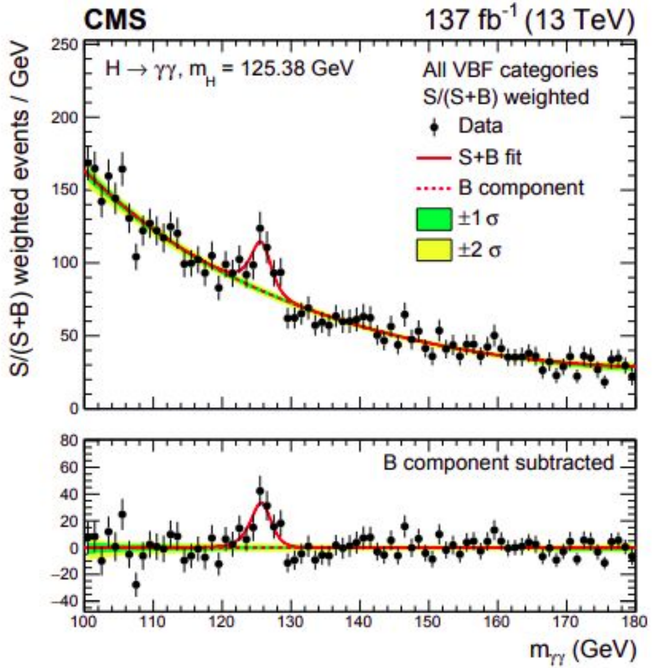
VBF $H \to \gamma\gamma$

JHEP 07 (2021) 027

→ Events are selected with two isolated photons.
→ High resolution of $m_{\gamma\gamma}$ [1-2% of m_H] has been utilized by fitting $m_{\gamma\gamma}$ distribution.
→ Signal is separated from ggF and non SM background by utilising di-photons and di-jets properties
→ Shape of the background derived directly from data using **discrete profile method**
→ UE & parton shower uncertainties is the dominant one for VBF mode



$$\mu_{VBF} = 1.04^{+0.14}_{-0.07}(\text{theo})^{+0.07}_{-0.06}(\text{exp})^{+0.31}_{-0.30}(\text{stat.})$$



VBF H→ WW

→ Analysis can probe direct **HVV coupling, V ~ W bosons**

→ Targets only the electrons and muons final states, with SF (eeνν/ μμνν) and DF (eμνν) final states

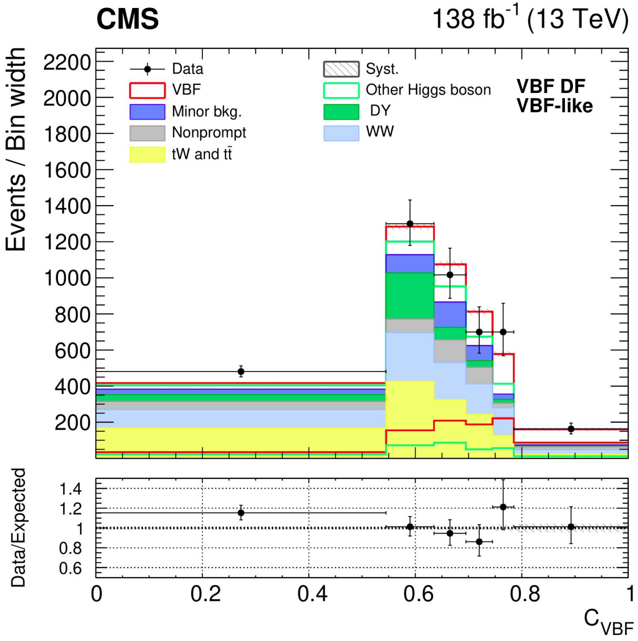
→ In DF, multi-classifier DNN used to distinguish VBF from ggF , WW and top backgrounds

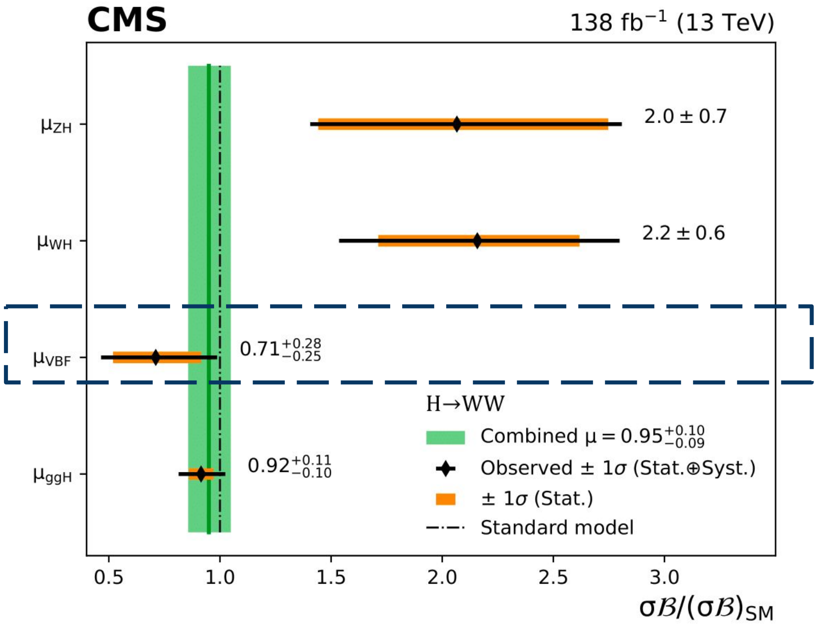
→ output of DNN used to fit

→ SF final states : DY → ll is the dominant source of background, dedicated MVA used to reduce it

→ Number of events used to fit

arXiv:2206.09466
Submitted to EPJC





Summary

→ VBF Higgs production has been explored in different Higgs decay modes in Run-2

→ The **combined signal strength of VBF H is 0.80 +/- 0.12** from Run-2 data, without inclusion of H→ bb

→ **VBF H→ bb has been presented for the first time from CMS using Run2 data**

→ The exclusive VBF H→ bb production is established with a significance of **2.4 σ (exp. = 2.7 σ)**

- the measured signal strength of the VBF production is $0.97^{+0.53}_{-0.45}$

→ Combination of VBF H→ bb with existing combined Run-2 VBF results would increase the sensitivity

→ Most of the measurements are still dominated by statistical uncertainties

→ All of the Higgs final states are represented in STXS framework as well

Recent STXS measurements from CMS – Aliya

→ Inclusion of Run-3 data will help to reduce it so please stay tuned for the Run-3 results and combination

Additional material

Detailed event selection VBF $H \rightarrow bb$

CMS PAS_HIG_22_009

Level	Requirements	2016 ($36.3\,\mathrm{fb}^{-1}$)		2018 ($54.5\,\mathrm{fb}^{-1}$)	
		TIGHT	LOOSE	TIGHT	LOOSE
HLT	p_T thresholds	$92, 76, 64, 16$ GeV		$105, 88, 76, 15$ GeV	
	number of b-tags	≥ 1	≥ 2	≥ 1	≥ 2
	$\Delta\phi_{bb}$	≤ 1.6	≤ 2.1	≤ 1.9	≤ 2.8
	$\Delta\eta_{jj}$	≥ 4.1	≥ 2.3	≥ 3.5	≥ 1.5
	m_{jj}	≥ 500 GeV	≥ 240 GeV	≥ 460 GeV	≥ 200 GeV
Offline	p_T thresholds	$95, 80, 65, 30$ GeV		$110, 90, 80, 30$ GeV	
	jet $\|\eta\| < 4.7$	✓	✓	✓	✓
	Lepton veto	✓	✓	✓	✓
	number of b-tags ≥ 2	✓	✓	✓	✓
	b-jet $\|\eta\| < 2.4$	✓	✓	✓	✓
	$\Delta\phi_{bb}$	≤ 1.6	≤ 2.1	≤ 1.6	≤ 2.1
	$\Delta\eta_{jj}$	≥ 4.2	≥ 2.5	≥ 3.8	≥ 2.5
	m_{jj}	≥ 500 GeV	≥ 250 GeV	≥ 500 GeV	≥ 250 GeV

ATLAS vs CMS [VBF $H \rightarrow bb$]

ATLAS : EPJC (2021) 81:537

CMS : PAS-HIG-22-009

Experiments	Inclusive sig. (σ)	VBF sig.(σ)	μ_{Hbb}	μ_{qqH}
ATLAS	2.7 (2.9)	2.6 (2.8)	0.95 +0.37/-0.35	0.95 +0.38/-0.36
CMS	2.5 (2.9)	2.4 (2.7)	0.92 +0.45/-0.39	0.97 +0.53/-0.45

VBF H$\rightarrow \tau\tau$

→ Why NN is more sensitive than CB

(i) Easy to distinguish qqH process than the ggH and Z $\rightarrow \tau\tau$
(ii) Reduce the migration of ggH events to the qqH category even after requiring 2 jets
(iii) Still qqH is statistically dominated whereas for ggH it is systematics dominated
(iv) In CB the systematic uncertainties appears for all individual variables used for the event selection, where as in NN the systematics is only on the 1D output score of theNN

Event selection : VBF H$\rightarrow \mu\mu$

JHEP 01 (2021) 148

Observable	VBF-SB	VBF-SR
Number of loose (medium) b-tagged jets Helps to reject tt bkgs.	$\leq$1 (0)	
Number of selected muons	=2	
Number of selected electrons	=0	
Jet multiplicity ($p_T > 25\,\text{GeV}$, $\lvert\eta\rvert < 4.7$)	$\geq$2	
Leading jet p_T	$\geq$35 GeV	
Dijet mass (m_{jj})	$\geq$400 GeV	
Pseudorapidity separation ($\lvert\Delta\eta_{jj}\rvert$)	$\geq$2.5	
Dimuon invariant mass	$110 < m_{\mu\mu} < 115\,\text{GeV}$ or $135 < m_{\mu\mu} < 150\,\text{GeV}$	$115 < m_{\mu\mu} < 135\,\text{GeV}$

Signal significance : VBF H$\rightarrow \mu\mu$

VBF is the most sensitive channel in H $\rightarrow \mu\mu$

JHEP 01 (2021) 148

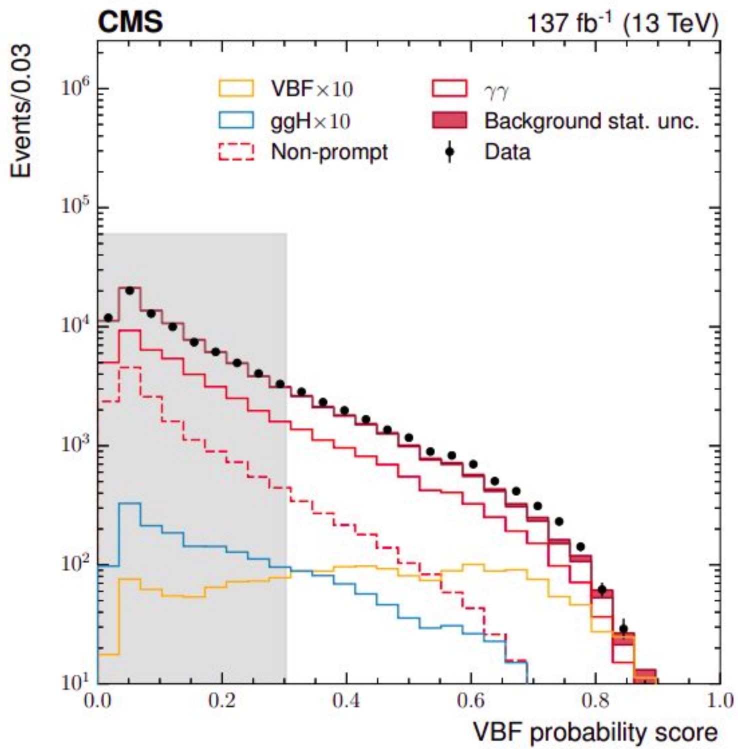
Production category	Observed (expected) signif.	Observed (expected) UL on μ
VBF	2.40 (1.77)	2.57 (1.22)
ggH	0.99 (1.56)	1.77 (1.28)
$t\bar{t}H$	1.20 (0.54)	6.48 (4.20)
VH	2.02 (0.42)	10.8 (5.13)
Combined $\sqrt{s} = 13\,\mathrm{TeV}$	2.95 (2.46)	1.94 (0.82)
Combined $\sqrt{s} = 7,\ 8,\ 13\,\mathrm{TeV}$	2.98 (2.48)	1.93 (0.81)

VBF H→ γγ

JHEP 07 (2021) 027

→ Leading (sub-leading) p_T > 40 (30) GeV
→ |η| < 4.7
→ m_{jj} > 350 GeV

→ Variables are mostly related to **di-jet** and di-photons
  → **Kinematics variables of jets**
  → **Angular variables of jets**
  → **Angular variables between photons and jets**
  → **p_T / $m_{\gamma\gamma}$**



VBF H$\rightarrow$ WW

arXiv:2206.09466
Submitted to EPJC

DF event selection

Subcategories	Selection
Global selection	
—	$p_{T1} > 25\,\text{GeV}$, $p_{T2} > 10\,\text{GeV}$ (2016) or $13\,\text{GeV}$ $p_T^{miss} > 20\,\text{GeV}$, $p_T^{\ell\ell} > 30\,\text{GeV}$, $m_{\ell\ell} > 12\,\text{GeV}$ $e\mu$ pair with opposite charge
2-jet VBF category	
SR	$60 < m_T^H < 125\,\text{GeV}$, $m_T(\ell 2, p_T^{miss}) > 30\,\text{GeV}$ 2 jets with $p_T > 30\,\text{GeV}$ no b-tagged jet with $p_T > 20\,\text{GeV}$ $m_{jj} > 120\,\text{GeV}$
Top quark CR	As signal region, no m_T^H requirement, $m_{\ell\ell} > 50\,\text{GeV}$ At least one b-tagged jet with $p_T > 30\,\text{GeV}$
$\tau\tau$ CR	As SR but with $m_T^H < 60\,\text{GeV}$ $30 < m_{\ell\ell} < 80\,\text{GeV}$

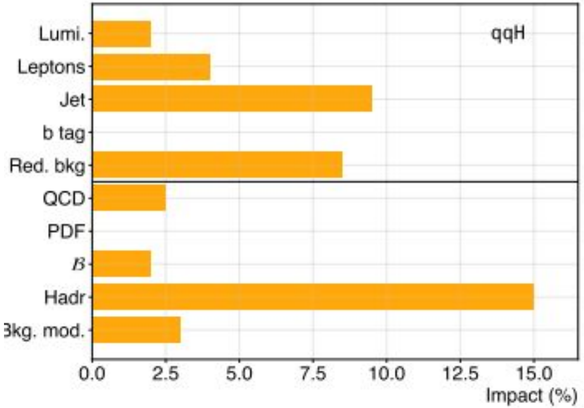
VBF H→ WW

arXiv:2206.09466
Submitted to EPJC

SF event selection

Subcategories	Selection
Global selection	
—	$p_{T1} > 25\,\mathrm{GeV}$, $p_{T2} > 10\,\mathrm{GeV}$ (2016) or $13\,\mathrm{GeV}$ $p_T^{miss} > 20\,\mathrm{GeV}$, $p_T^{\ell\ell} > 30\,\mathrm{GeV}$ $m_{\ell\ell} > 12\,\mathrm{GeV}$, $\lvert m_{\ell\ell} - m_Z \rvert > 15\,\mathrm{GeV}$ ee or $\mu\mu$ pair with opposite charge No b-tagged jets with $p_T > 20\,\mathrm{GeV}$
2-jet VBF category	
ee, $\mu\mu$	$m_{\ell\ell} < 60\,\mathrm{GeV}$, $65 < m_T^H < 150\,\mathrm{GeV}$ At least 2 jets with $p_T > 30\,\mathrm{GeV}$ $\lvert \Delta\phi_{\ell\ell} \rvert < 1.6$, $m_{jj} > 350\,\mathrm{GeV}$ DYMVA above threshold
WW CR	As signal region, $m_{\ell\ell} > 100\,\mathrm{GeV}$ $m_T^H > 60\,\mathrm{GeV}$, $m_T(\ell 2, p_T^{miss}) > 30\,\mathrm{GeV}$
Top quark CR	As signal region, $m_{\ell\ell} > 100\,\mathrm{GeV}$, $m_T(\ell 2, p_T^{miss}) > 30\,\mathrm{GeV}$ At least one of the leading jets b-tagged

VBF H→ ZZ



	Expected	Observed
$\mu_{\mathrm{t\bar{t}H,tH}}$	$1.00^{+1.23}_{-0.77}$ (stat) $^{+0.51}_{-0.06}$ (syst)	$0.17^{+0.88}_{-0.17}$ (stat) $^{+0.42}_{-0.00}$ (syst)
μ_{WH}	$1.00^{+1.83}_{-1.00}$ (stat) $^{+0.75}_{-0.00}$ (syst)	$1.66^{+1.52}_{-1.66}$ (stat) $^{+0.85}_{-0.00}$ (syst)
μ_{ZH}	$1.00^{+4.79}_{-1.00}$ (stat) $^{+6.76}_{-0.00}$ (syst)	$0.00^{+4.38}_{-0.00}$ (stat) $^{+3.24}_{-0.00}$ (syst)
μ_{VBF}	$1.00^{+0.53}_{-0.44}$ (stat) $^{+0.18}_{-0.12}$ (syst)	$0.48^{+0.46}_{-0.37}$ (stat) $^{+0.14}_{-0.10}$ (syst)
$\mu_{\mathrm{ggH,b\bar{b}H}}$	1.00 ± 0.10 (stat) $^{+0.12}_{-0.10}$ (syst)	0.99 ± 0.09 (stat) $^{+0.11}_{-0.09}$ (syst)
μ	$1.00^{+0.08}_{-0.07}$ (stat) $^{+0.10}_{-0.08}$ (syst)	0.94 ± 0.07 (stat) $^{+0.09}_{-0.08}$ (syst)

	$(\sigma\mathcal{B})_{\mathrm{obs}}$ (fb)	$(\sigma\mathcal{B})_{\mathrm{SM}}$ (fb)	$(\sigma\mathcal{B})_{\mathrm{obs}}/(\sigma\mathcal{B})_{\mathrm{SM}}$
ttH	3^{+16}_{-3}	15.9 ± 1.4	$0.16^{+0.98}_{-0.16}$
VH-lep	41^{+52}_{-35}	25.9 ± 0.8	$1.56^{+1.99}_{-1.34}$
qqH	61^{+53}_{-44}	122 ± 6	$0.50^{+0.44}_{-0.36}$
ggH	1214^{+135}_{-125}	1192 ± 95	$1.02^{+0.11}_{-0.10}$
Inclusive	1318^{+130}_{-122}	1369 ± 164	$0.96^{+0.10}_{-0.09}$