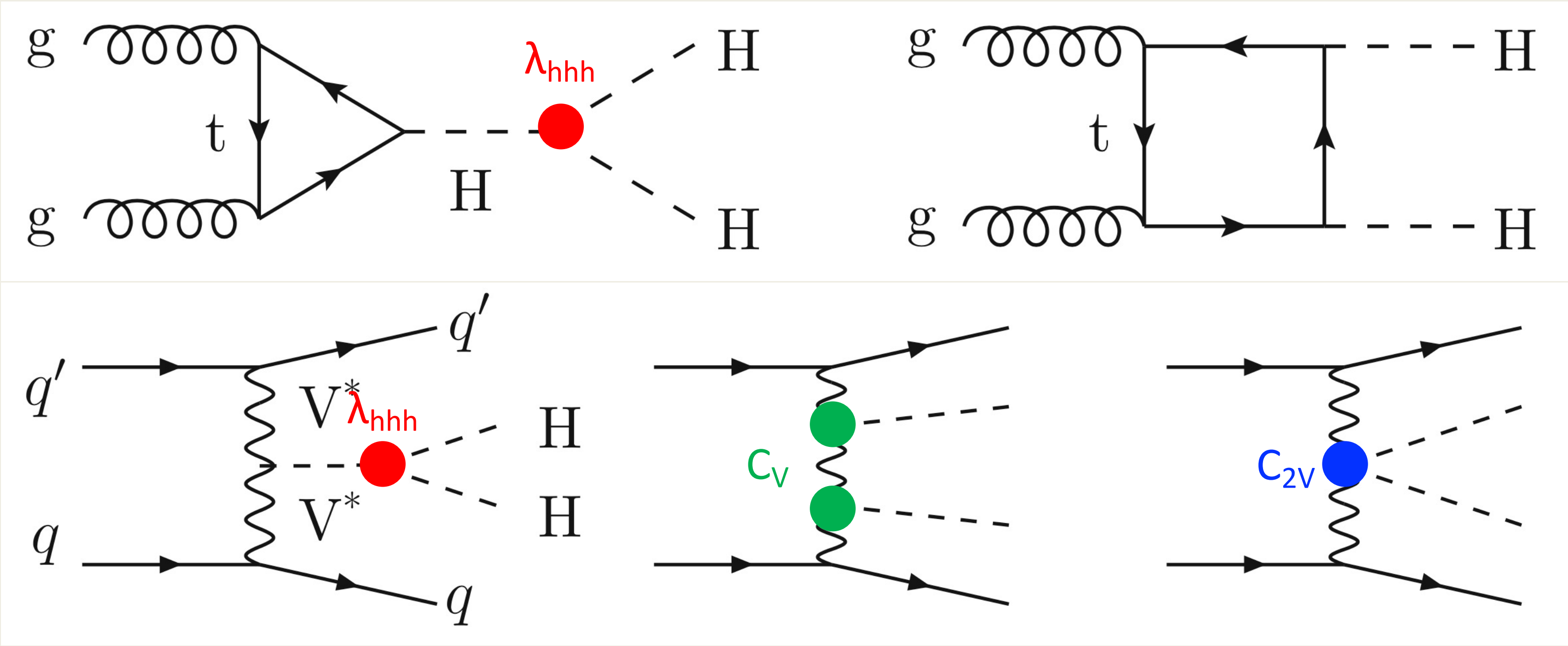
DOUBLE HIGGS SEARCHES AT CMS

Davide Zuolo
INFN Milano - Bicocca
On behalf of the CMS collaboration

Standard Model (SM)

- The measurement of $\sigma(HH)$ is the best way to extract the Higgs self-coupling λ_{HHH}
- This parameter defines the shape of the Higgs potential together with m_H and the vacuum expectation value (v)

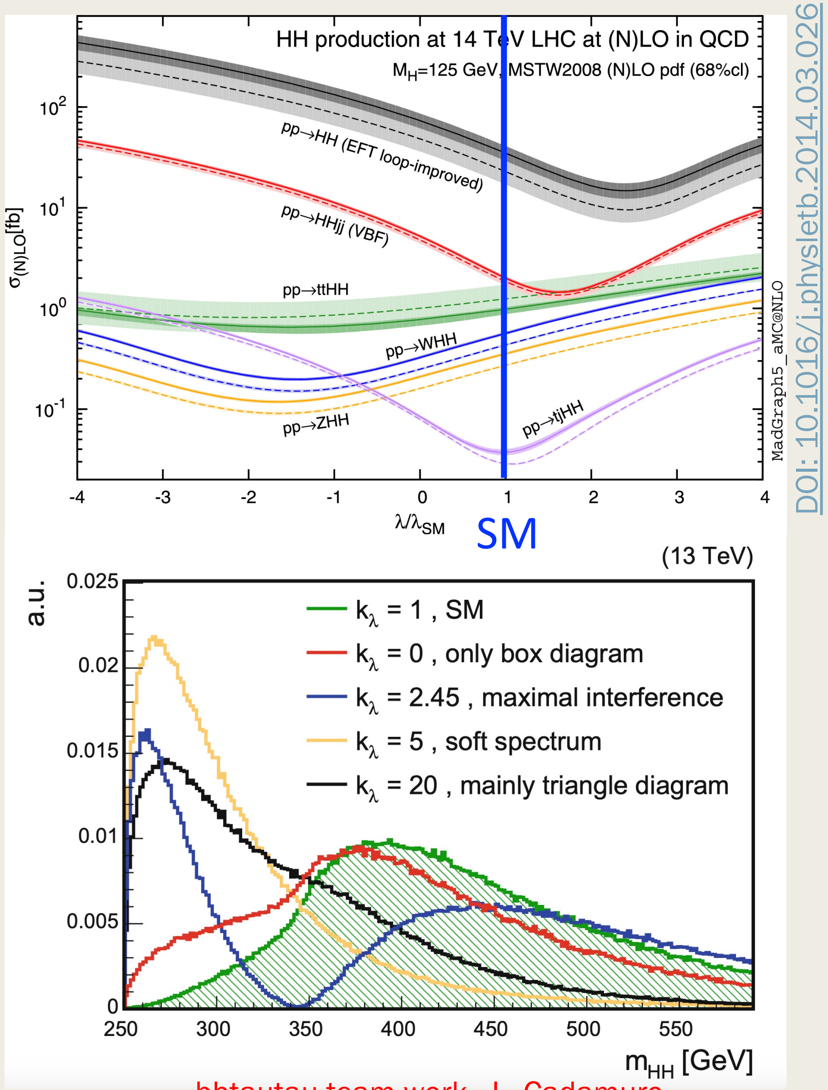


Gluon-Gluon Fusion (GGF): Destructive interference between the two diagrams leads to a small cross section (31.05 fb)

Vector Boson Fusion (VBF): Cross section: 1.726 fb
Unique handle to probe VVHH coupling (C_{2V})

Beyond Standard Model (BSM)

- Anomalous values of the couplings affects both the cross section and the kinematics of HH production.
- They are modelled through coupling modifiers
  - $k_\lambda = \lambda_{HHH}/\lambda_{HHH,SM}$
  - $k_{2V} = C_{2V}/C_{2V,SM}$
  - $k_V = C_V/C_{V,SM}$
- Searches for X → HH → xxyy are motivated by models with a warped extra dimension (WED), as proposed by Randall and Sundrum
  - *spin-0 radion [1-2-3]*
  - *spin-2 first Kaluza–Klein (KK) excitation of the graviton [4-5-6]*



bbtautau team work - L. Cadamuro

Outline

Analyses made public in 2021 using the full LHC Run II (2016+2017+2018) dataset (138 fb^{-1})

- Non-resonant analyses
  - GGF + VBF HH → bbbb (CMS-PAS-HIG-20-005)
  - VBF HH → bbbb boosted (CMS-PAS-B2G-21-001)
- Resonant analyses
  - X → HH → bb+leptons boosted (CMS-PAS-B2G-20-007)
  - X → HH → bbbb boosted (CMS-PAS-B2G-20-004)
  - X → YH → bbtautau (CMS-PAS-HIG-20-014, DOI)
  - X → YH → bbbb boosted (CMS-PAS-B2G-21-003)

  } See talk by Devdatta Majumder
- Non-resonant + resonant analysis: HH → MultiLepton (CMS-PAS-HIG-21-002)

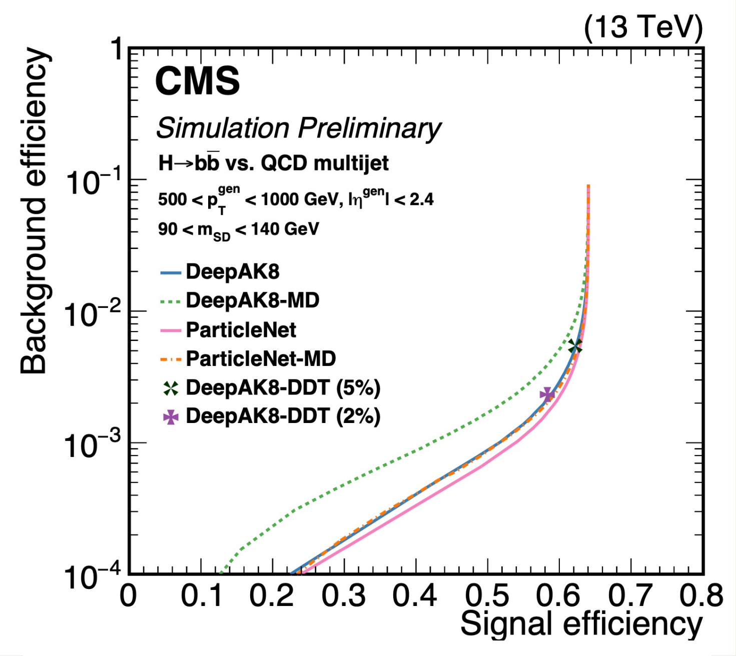
bbbb final state (highest BR)

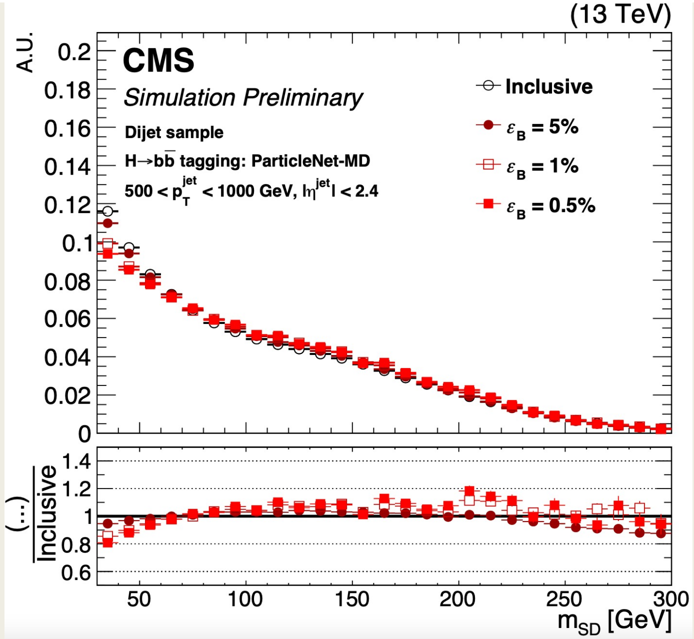
Resolved

- Reconstruct H → bb decays from AK4 (small radius) jets
- b-jets tagged using DNN-based deepJet algorithm
- 4 b-jets → three possible pairings → large combinatorial
- "GGF killer" BDT to separate GGF and VBF events

Boosted

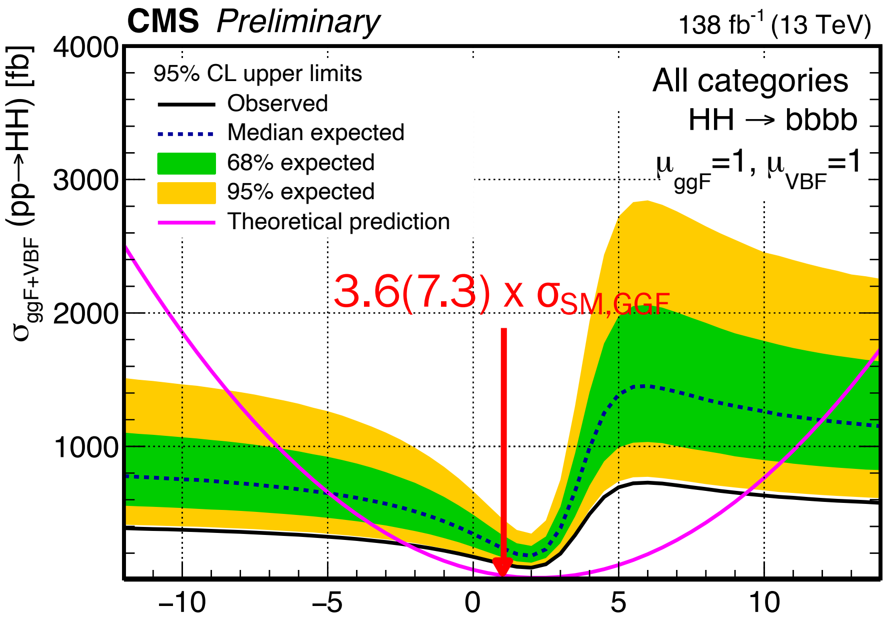
- Reconstruct H → bb decays as AK8 (large radius) jets
- Use AK8 jet substructure to identify H→bb decays with ParticleNet algorithm [Graph Neural Network - 1,2]
- MD = "Mass Decorrelated" → No sign of sculpting of jet mass (crucial for background estimate!)
- Take the two highest p_T AK8 jets as H_1 and H_2



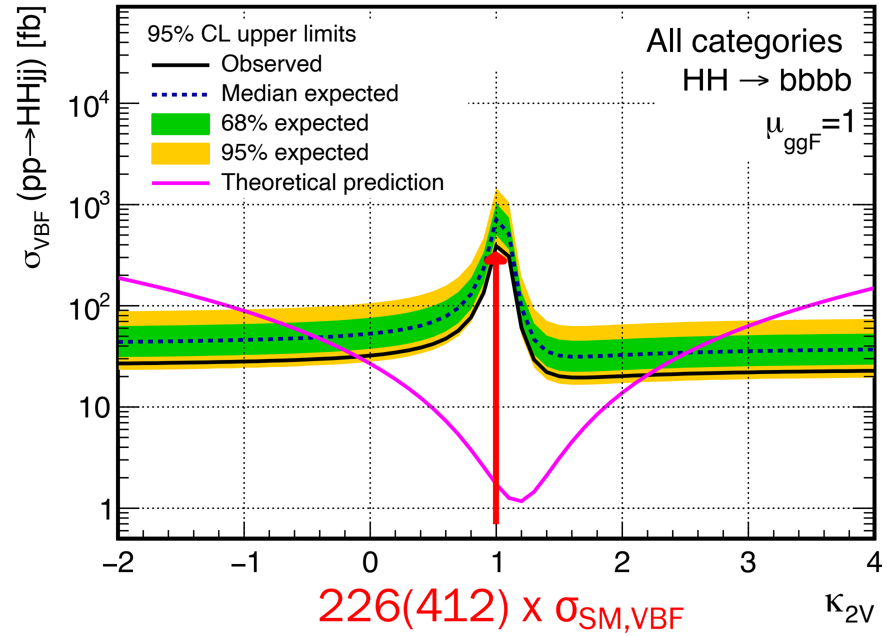


GGF + VBF HH → bbbb

- b-jet pairs ordered according to $D = \frac{|m_{H1} - k m_{H2}|}{1+k^2}$ → correct jet pairing ranging between 82 – 96% (91 – 98%) for the different couplings in ggF (VBF) signal events
- events categorized according to GGF-killer BDT score → GGF, VBF-SM and VBF-BSM enriched categories
- large contamination from QCD multijet production and $t\bar{t}$ +jets → Fully data driven background estimation
- Most stringent constraint on $\sigma_{SM,GGF}$ achieved so far



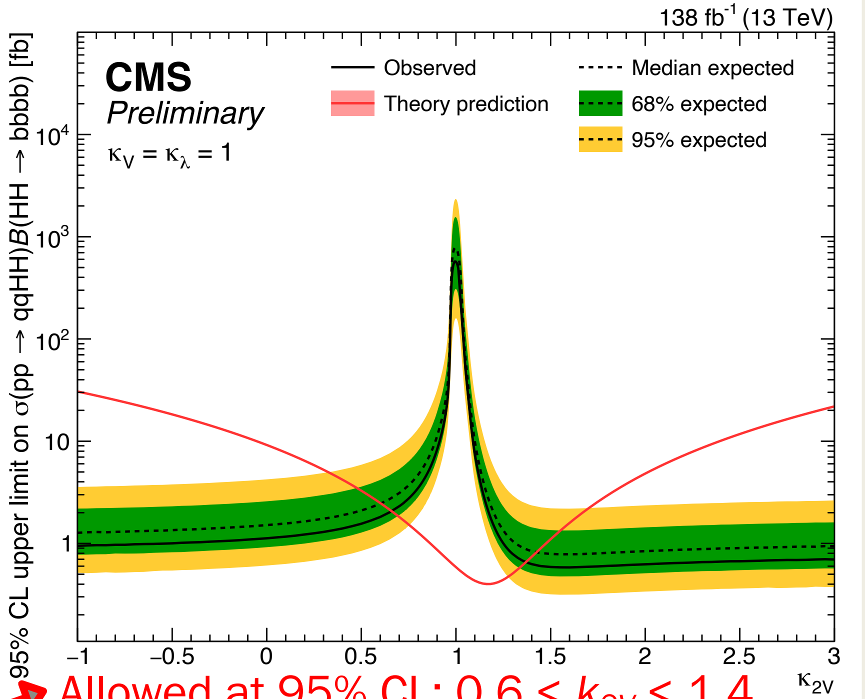
Allowed at 95% CL: $-2.3 < k_\lambda < 9.4$



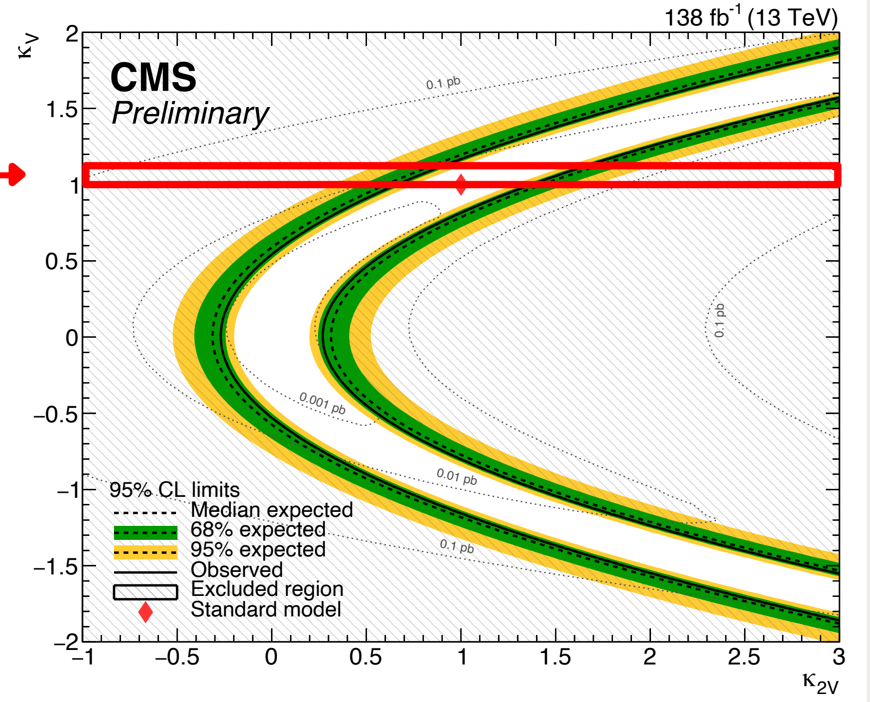
Allowed at 95% CL: $-0.1 < k_{2V} < 2.2$

VBF HH → bbbb boosted

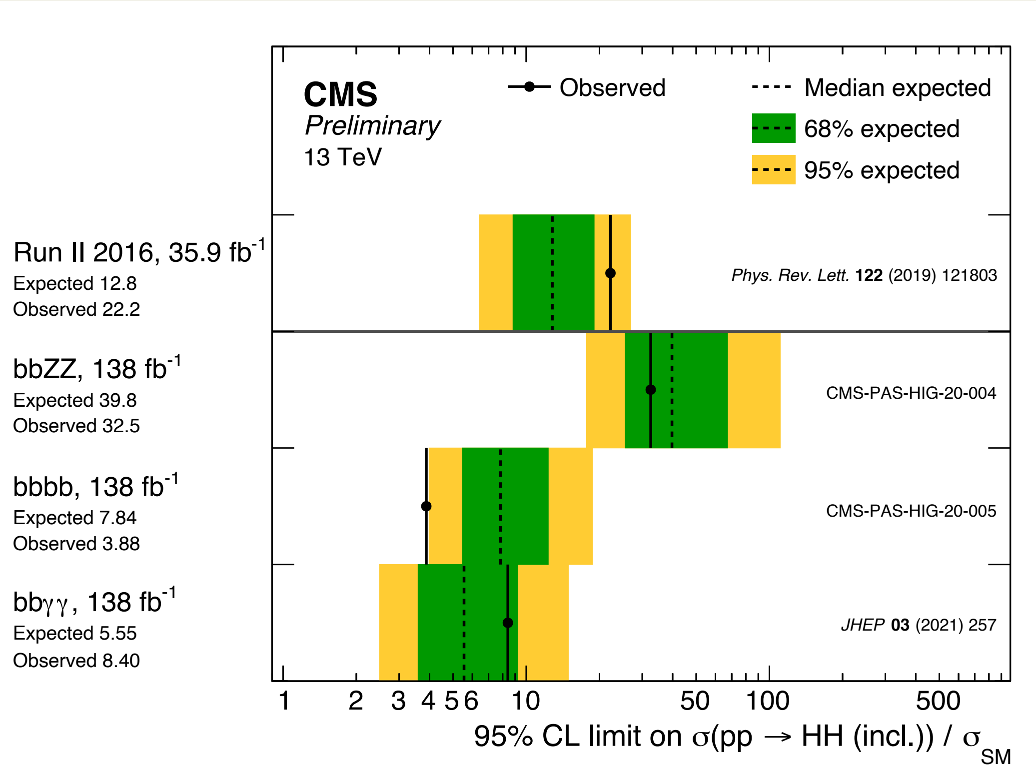
- In BSM scenarios with modified couplings, a significant fraction of signal becomes boosted → Enhanced sensitivity for anomalous couplings
- Reconstruct Higgs candidate mass with ParticleNet-based regression algorithm
- Use m_{HH} as observable in three purity categories
- Strongest constraint on k_{2V} achieved so far
- When the external constraints on k_V (from VH processes measurements) are included, $C_{2V} = 0$ is strongly disfavored



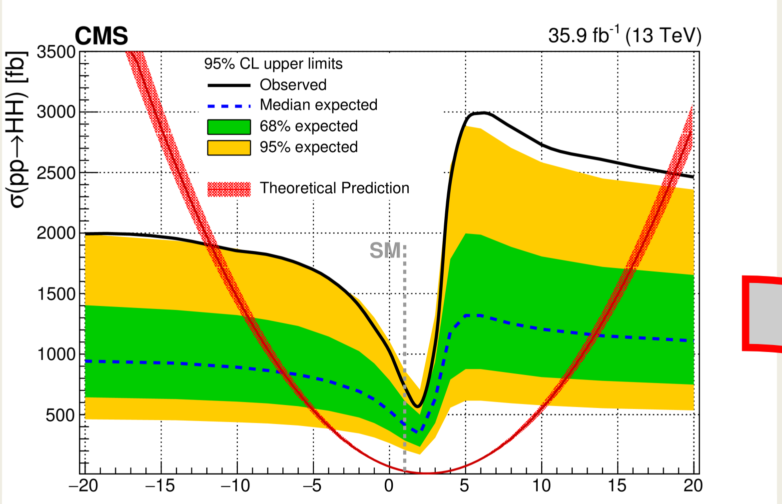
Allowed at 95% CL: $0.6 < k_{2V} < 1.4$



Non-resonant analyses summary

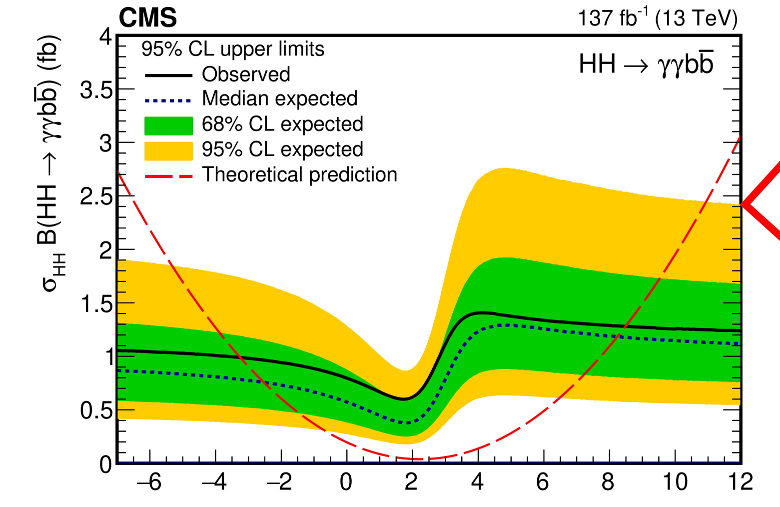


Each of the expected limits of 4b and bb$\gamma\gamma$ with run 2 statistics are 2 times more stringent than the 2016 combination!



Allowed at 95% CL: -11.8< k_λ < 18.8

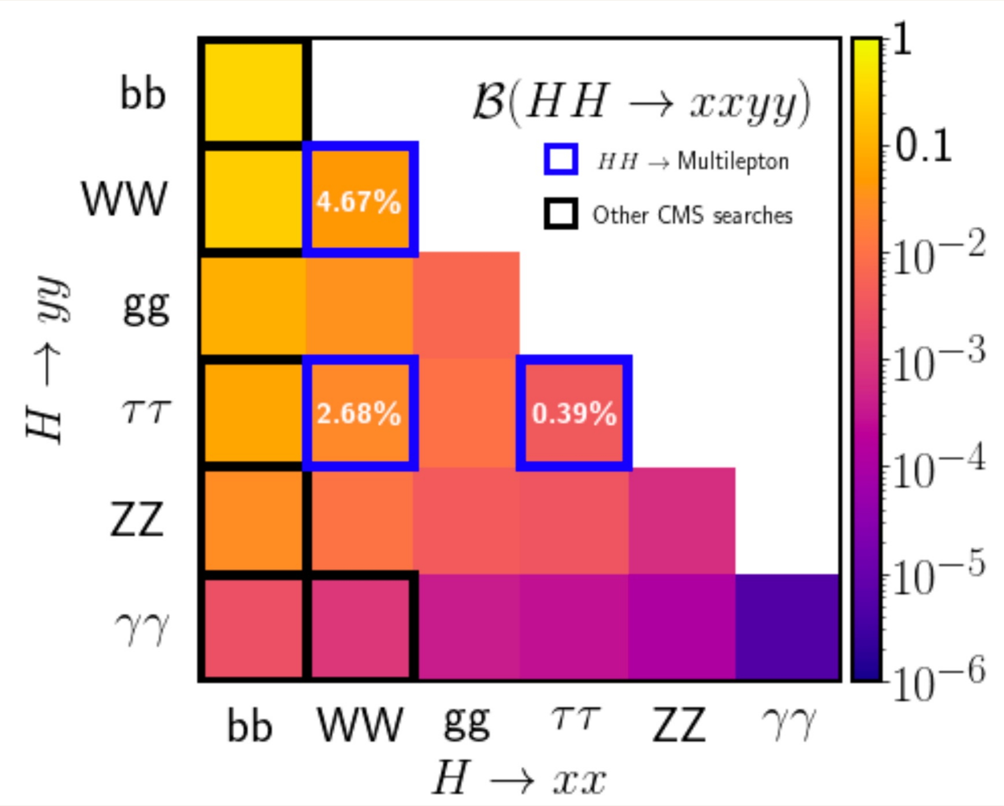
2016 comb vs $\gamma\gamma$bb Full Run 2



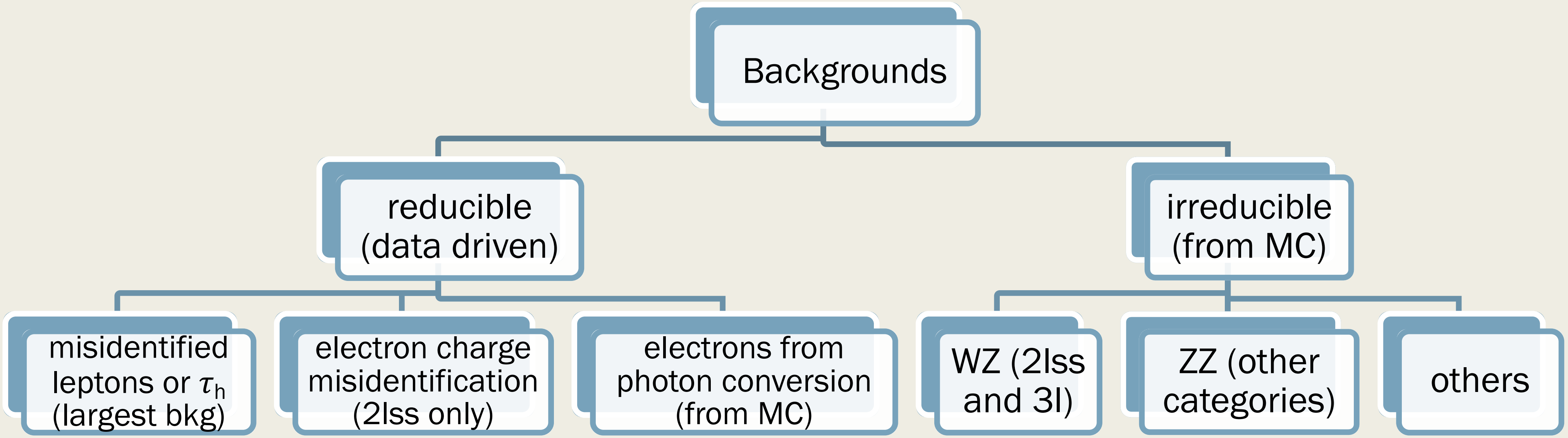
Allowed at 95% CL: -3.3 < k_λ < 8.5

NEW: HH → MultiLepton - 1

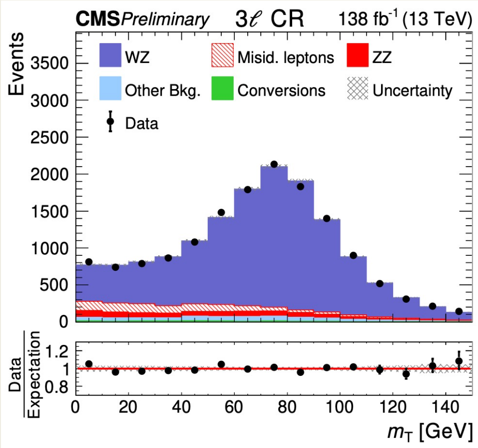
- Target decay modes: HH → WWWW / WW$\tau\tau$ / $\tau\tau\tau\tau$, covering ~ 7.7% of the HH decays
- 7 search categories, distinguished by the number of reconstructed electrons or muons (l) and tau decaying to hadrons (τ_h)
  - *4l* , *3l+0τ_h*, *2lss+0/1τ_h*,
  - *3l+1τ_h*, *1l+3τ_h*, *2l+2τ_h* , *0l+4τ_h*.
- Various hypotheses tested:
  - *SM prediction*
  - *anomalous values of SM couplings*
  - *Effective Field Theory interpretation (not in this talk)*
  - *resonant production (spin-0 and spin-2 resonances with masses between 250 and 1000 GeV)*

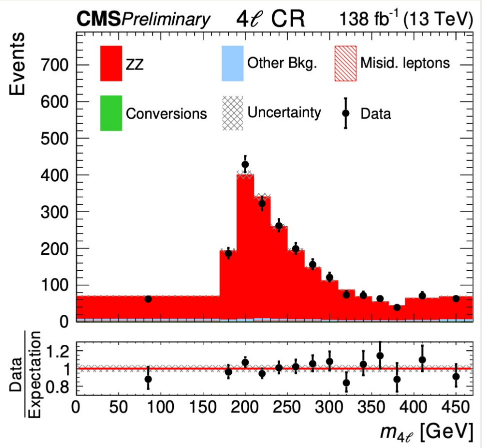


NEW: HH → MultiLepton - 2



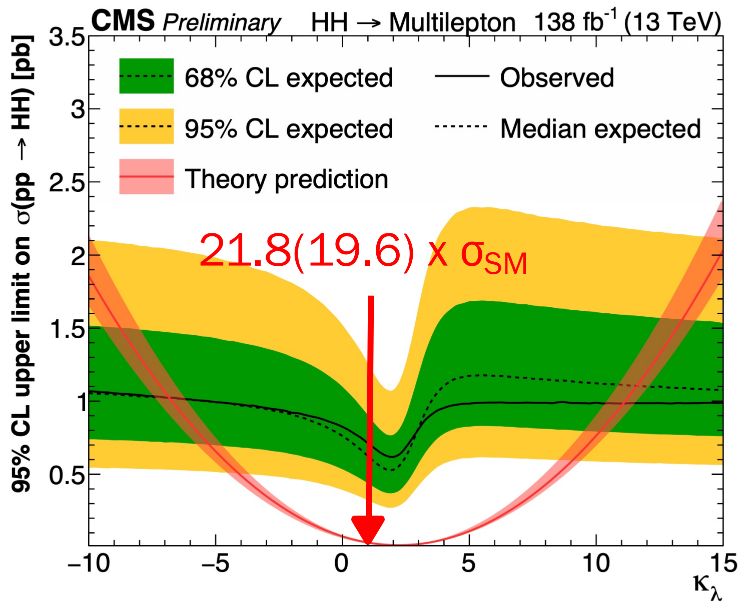
Dedicated BDT trained to separate prompt leptons (from W, Z and τ decay) from nonprompt (from hadron decay) or misidentified leptons.



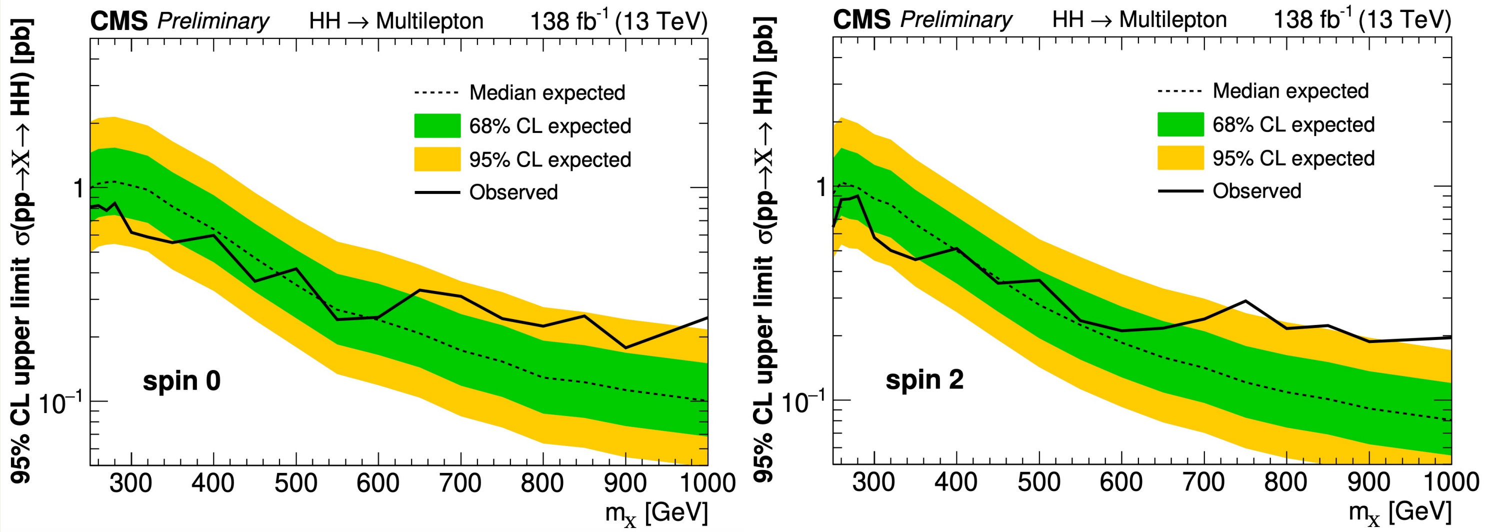


NEW: HH → MultiLepton - 3

- Three BDT classifiers trained for each of the seven search categories:
  - *nonresonant HH production vs bkg (Higgs couplings as inputs)*
  - *spin-0 and spin-2 resonance vs bkg (resonance mass as input)*
- 3l WZ and 4l ZZ control regions included in the fit



Allowed at 95% CL: $-7.0 < k_\lambda < 11.2$



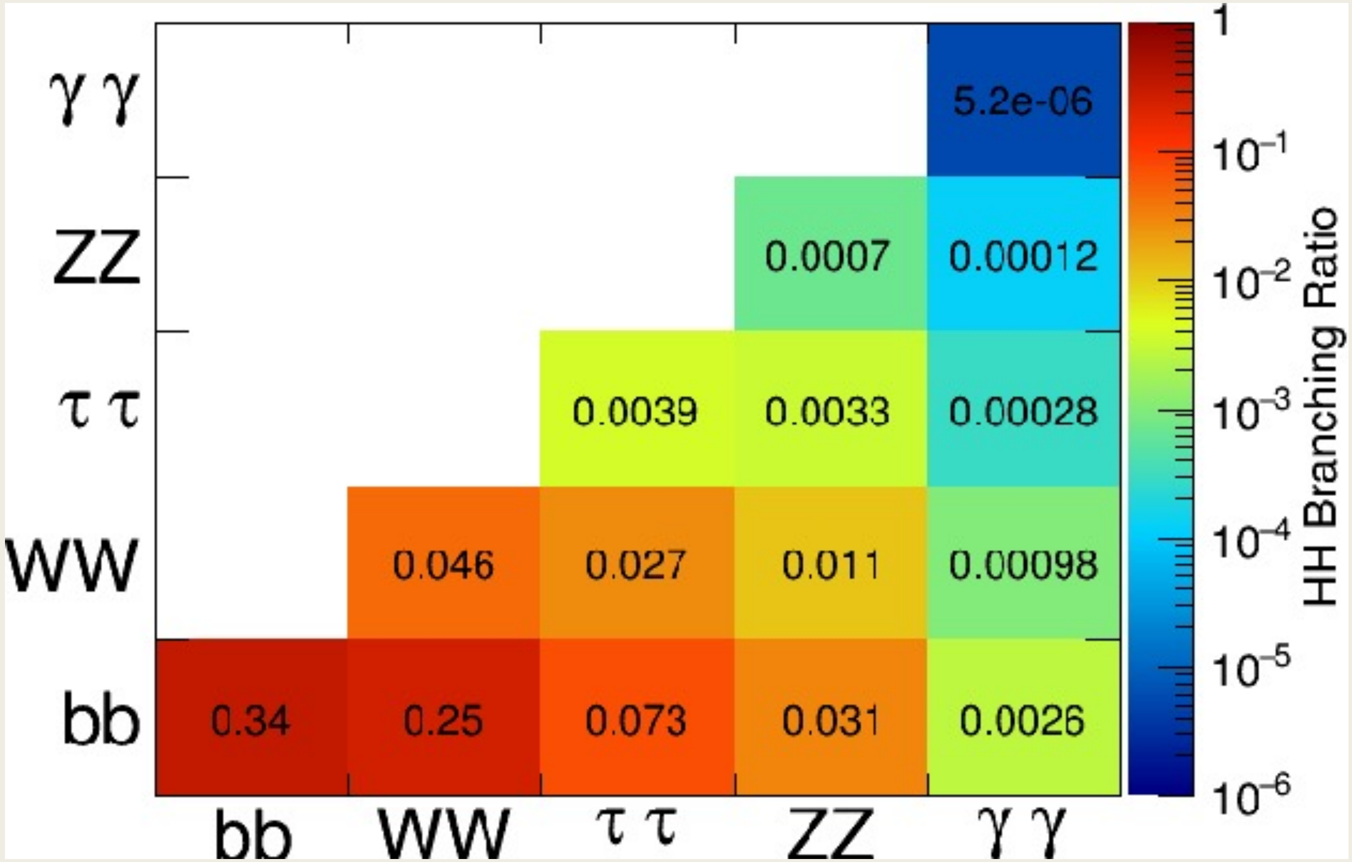
Small excess (2.1 std dev significance) of data event near m_X = 750 GeV in the 2lss and 3l categories affecting observed limits at high mass

Conclusions

- Searches for HH production using the LHC Run 2 dataset have run at full swing in CMS during 2020 and 2021
- Two of the three most sensitive non-resonant searches (bbbb and bb$\gamma\gamma$) have been made public → each of them sees factor 2 improvement compared to 2016 combination
- Combination of HH non resonant searches (including the third most sensitive channel, bb$\tau\tau$) foreseen for spring 2022
- Many interesting results from the investigation of boosted final states
- Resonant searches (both X → HH and X → YH) are being performed and a combination of them is foreseen in the second half of 2022

Additional material

Branching Ratios



Previous H→bb ID algorithms (PAS)

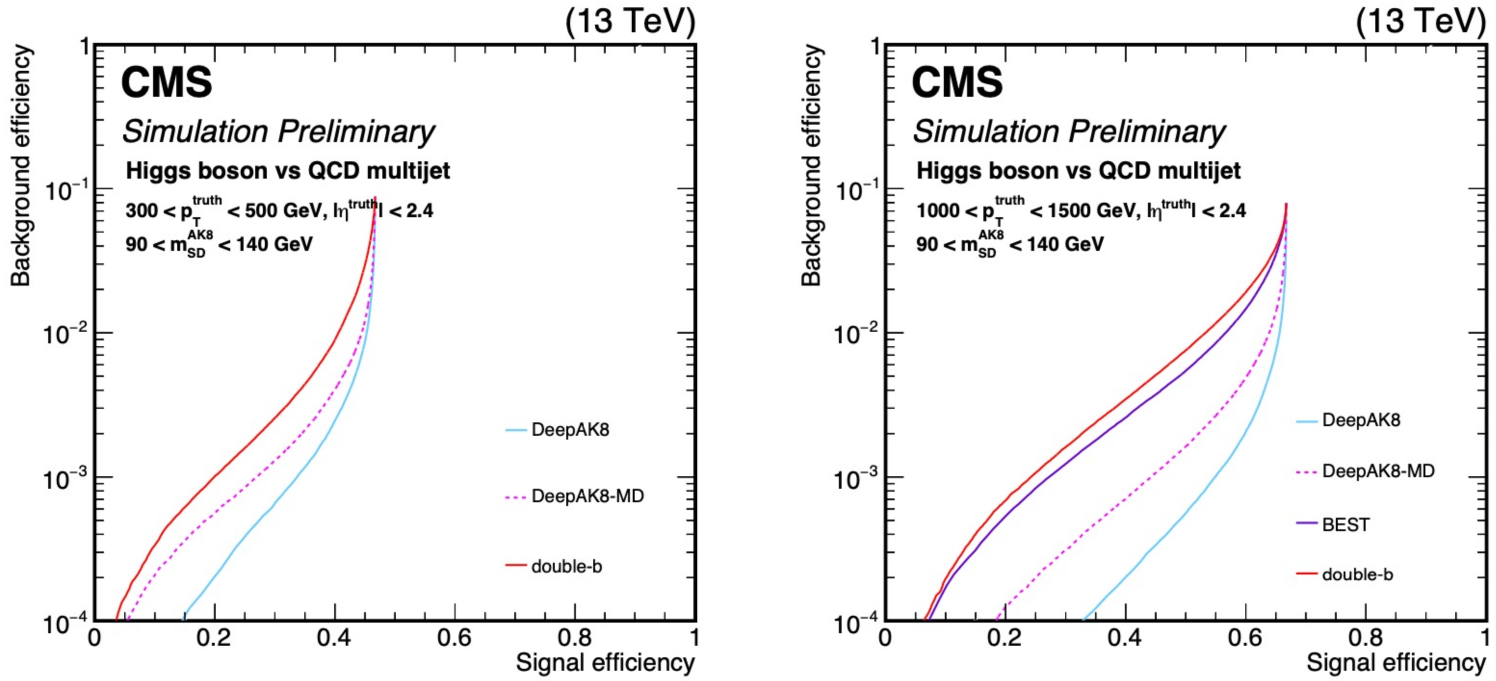


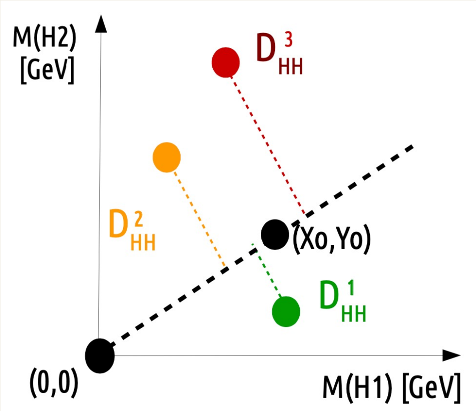
Figure 14: Performance comparison of the hadronically decaying H boson identification algorithms in terms of receiver operating characteristic (ROC) curves in two regions based on the p_T of the truth particle; Left: $300 < p_T < 500$ GeV, and Right: $1000 < p_T < 1500$ GeV. The H boson is forced to decay in a pair of b quarks. Additional fiducial selection criteria applied to the jets are displayed on the plots.

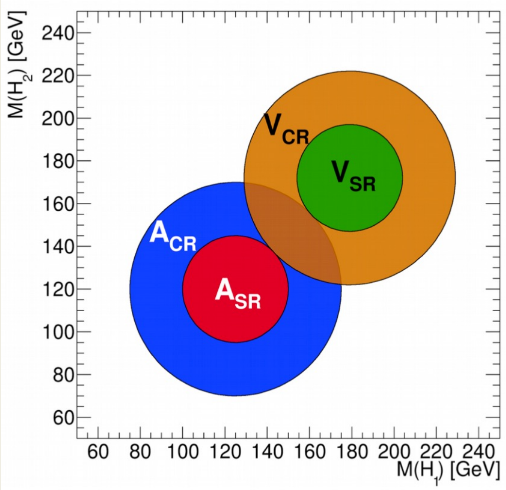
GGF HH → bbbb - 1

- Trigger selections require the presence of at least four jets, three b-tagged jets, and a minimal jet momentum scalar sum (H_T)
- b-jet candidates: Pt > 30-40 GeV, central, deepJet medium WP, PUJetID, PFJetID
- pairings ordered according to $D = \frac{|m_{H1} - k m_{H2}|}{1+k^2}$ → correct jet pairing ranging between 82 – 96% (91 – 98%) for the different couplings in ggF (VBF) signal events
- Reject events with one reconstructed muon or electron
- VBF-jet candidates (excluding b-jets) $P_T >$ 25 GeV, $|\eta| < 4.7$, PUJetID, PFJetID
- VBF-jet pair selection: Two highest P_T jets with $\eta(j_1) \times \eta(j_2) < 0$

Background estimation is fully data-driven

- Analysis (A) and Validation (V) regions
- '4b' and '3b' region (4th jet btag score fails medium WP)
- BDT-based reweighting to correct for differences between the 3b and the 4b regions





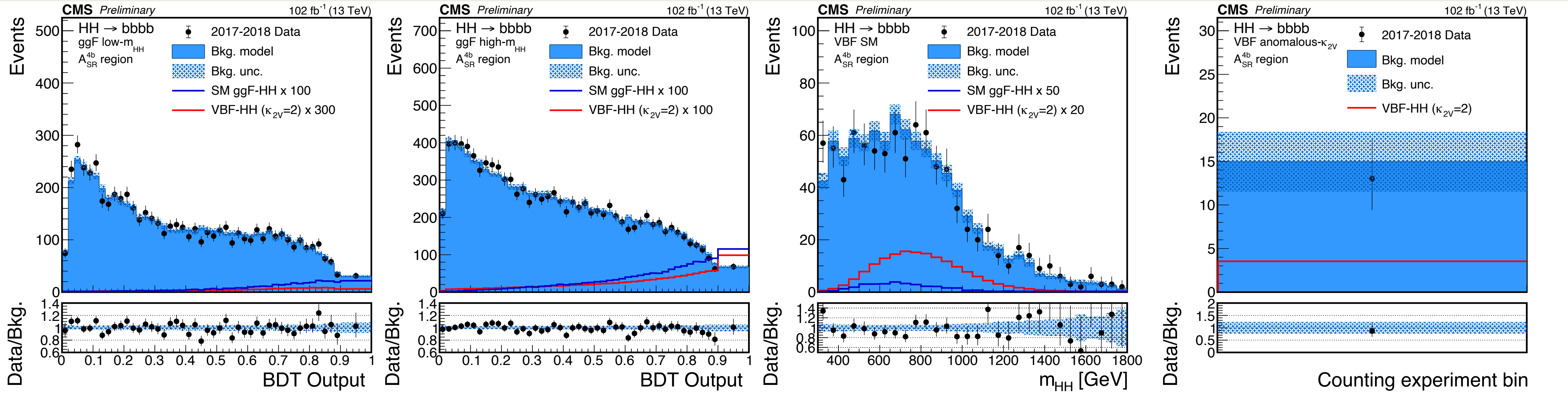
GGF + VBF HH → bbbb - 2

<u>Gluon fusion categorization</u>

- ggF cat1 or Low-m_{HH}: $m_{HH} < 450$ GeV
- ggF cat2 or High-m_{HH}: $m_{HH} \geq 450$ GeV
- Observable: BDT output distribution
  - *SM ggF vs data-driven bkg model*
  - *Trained by category*

<u>Vector Boson Fusion categorization</u>

- VBF cat1 or VBF-SM:
  - *$0.5 \leq ggfKiller < 0.97$*
  - *Observable: m_{HH} distribution*
- VBF cat2 or VBF-anomalous-k_{2V}:
  - *$0.97 \leq ggfKiller \leq 1.0$*
  - *Observable: Counting experiment*



VBF HH → bbbb boosted - 1

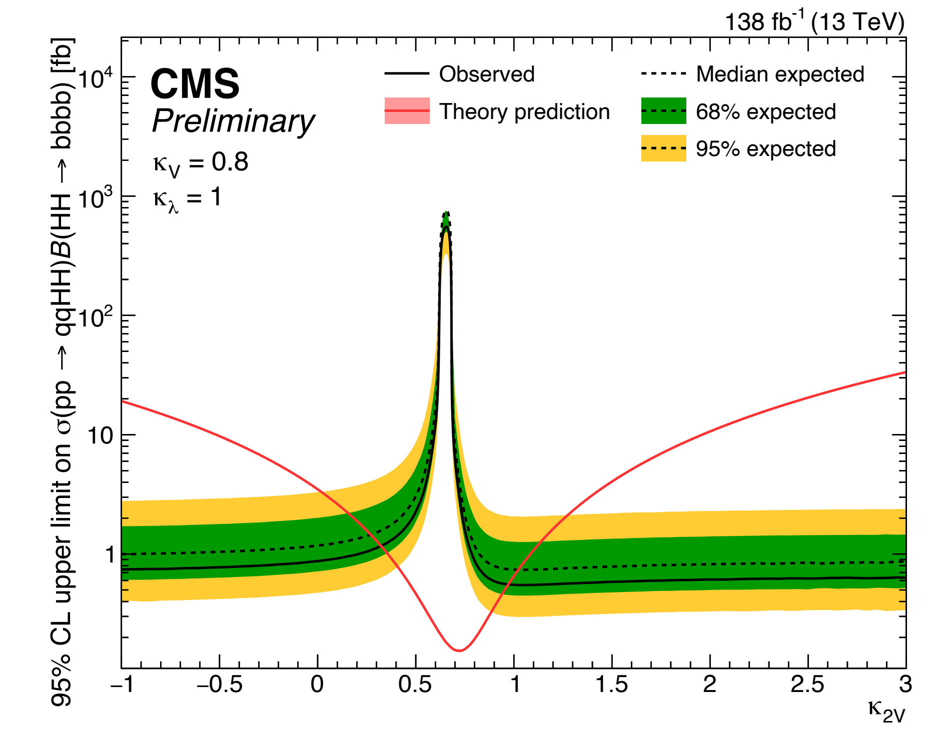
Trigger	Combination of H_T and single-jet triggers
Higgs boson candidates	≥ 2 large-radius jets with $\lvert\eta\rvert < 2.4$ $p_T^{lead} > 500\,GeV$, $p_T^{subl} > 400\,GeV$ $\Delta\phi > 2.6$, $\Delta\eta < 2.0$
Lepton veto	$N_e = 0$, $N_\mu = 0$
$H \to b\bar{b}$ identification with ParticleNet	Three exclusive search categories based on D_{bb} working points: high purity (HP), medium purity (MP) and low purity (LP)
VBF selections	≥ 2 small-radius jets with $p_T > 25\,GeV$, $\lvert\eta\rvert < 4.7$ $m_{jj} > 500\,GeV$, $\Delta\eta_{jj} > 4.0$
Signal mass range	$110 < m^{lead} < 150\,GeV$, $100 < m^{subl} < 145\,GeV$

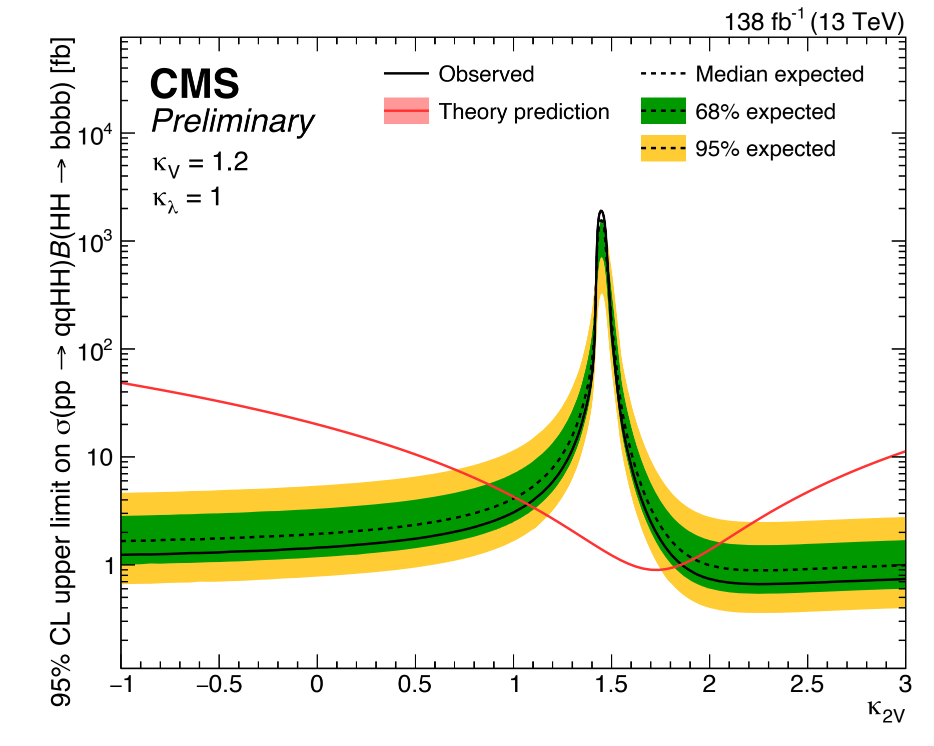
$$D_{bb} = \frac{P(X \to bb)}{P(X \to bb) + P(QCD)}$$

Background contamination very limited, tail of SM processes

- **TTbar** background from simulation, with corrections from a top-enriched region
- **QCD multijet** background estimated with a data-driven method (ABCD)

VBF HH → bbbb boosted - 2

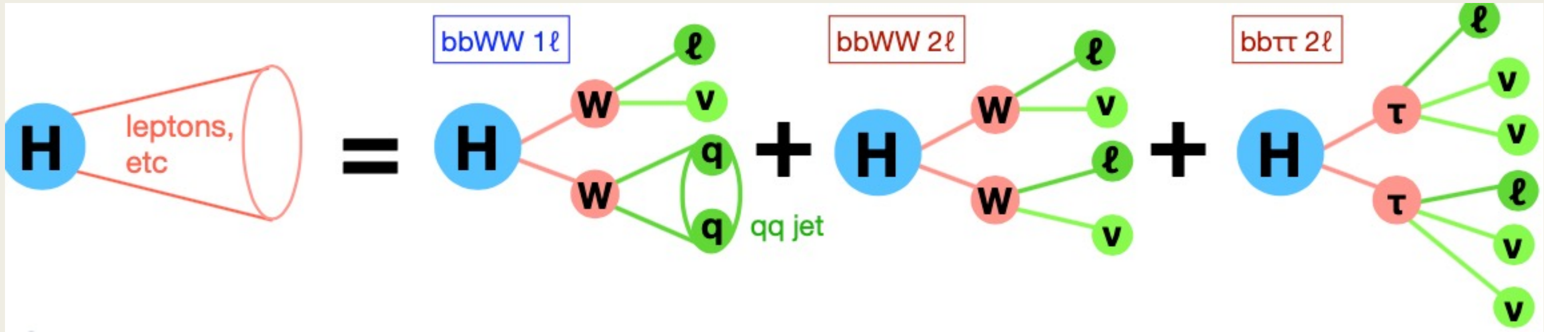




While the excluded k_{2V} range changes depending on the value of k_V, the $k_{2V} = 0$ hypothesis remains excluded also in these scenarios with varied k_V values.

X → HH → bb+leptons boosted - 1

- X decays to two boosted Higgs bosons, ~ back-to-back in the boosted regime
- One Higgs decays via H⟶bb → identified as large radius jet using [DeepAK8]()
  - *[Soft-drop]() mass m_{bb} of the jet is one of two search variables*
- Other Higgs decays into a final state with leptons (H⟶WW* and H⟶ $\tau\tau$)
  - ***Single-lepton (1l) channel***
    - Reconstruct W⟶qq as large radius jet → No mass constraint since W can be on- or off-shell
    - Lepton is often very close to the qq jet → Loose IDs and isolation, tailored for leptons in jets
  - ***Dilepton (2l) channel***
    - No jet ⟶ cleaner event than in 1l
    - Kinematic reconstruction of the leptonic H boson four momentum is possible
- Invariant mass of H⟶bb and leptonic H boson (m_{HH}) is second search variable

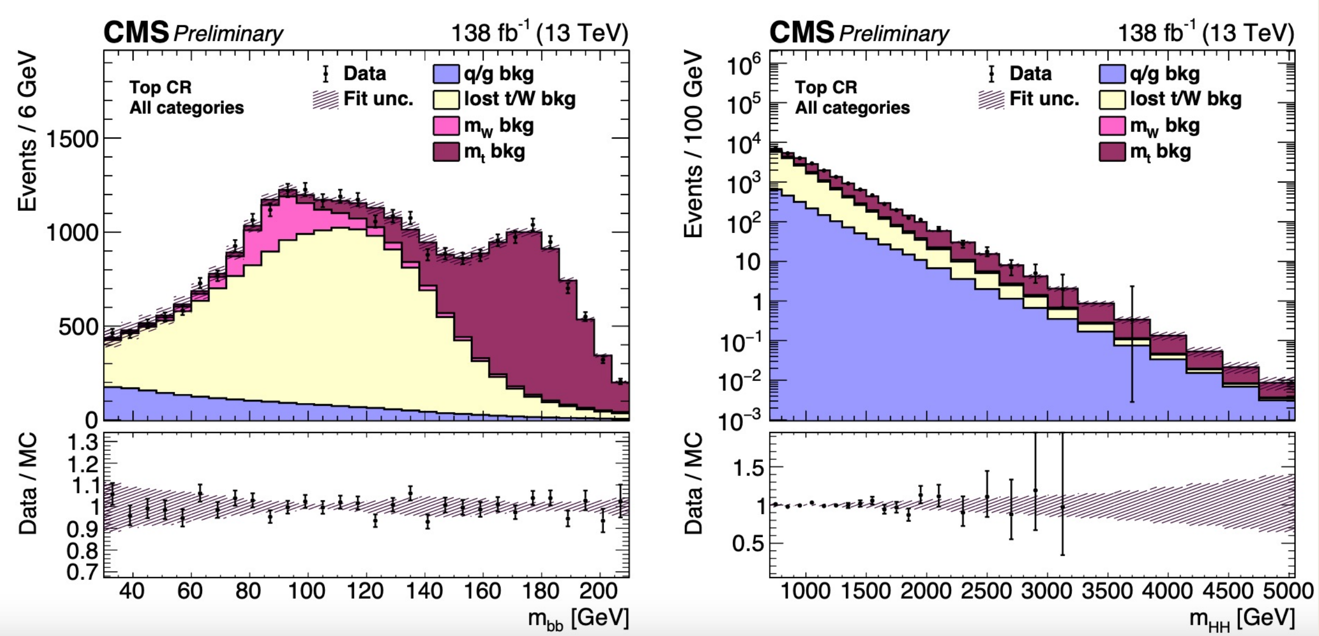


X → HH → bb+leptons boosted - 2

- Background events are split in four components based on the number of generator-level quarks from the immediate decay of a top quark or vector boson within $\Delta R < 0.8$ of the bb jet axis.

Bkg. category	Dominant SM processes	Resonant in $m_{b\bar{b}}$	Num. of gen-level quarks
m_t	$t\bar{t}$	top mass	3 from top
m_W	$t\bar{t}$	W mass	2 from W
lost t/W	$t\bar{t}$	No	1 or 2
q/g	V+jets and multijet	No	0

- Two step production of templates in the m_{bb} and m_{HH} mass plane
  1. *produce inclusive templates combining events in multiple categories with relaxed selections*
  2. *fit the inclusive templates to the simulated mass distributions for each of the 12 event categories to produce a unique template.*



X → HH → bb+leptons boosted - 3

- Final fit performed 12 categories
  - *2 same flavour (ee + $\mu\mu$)*
  - *2 opposite flavour (eμ)*
  - *4 single electron*
  - *4 single muon*
- Dilepton categories more sensitive than single lepton
- 6x (14x) improvement at low(high) m_X w.r.t. 2016 only analysis (single lepton only)