Search for nonresonant Higgs boson pair production in $b\bar{b}\tau^+\tau^-$ final state in pp collisions at $\sqrt{s} = 13$ TeV

Final Exame.

Matteo Pisano — 13th July 2023

The process
Introduction

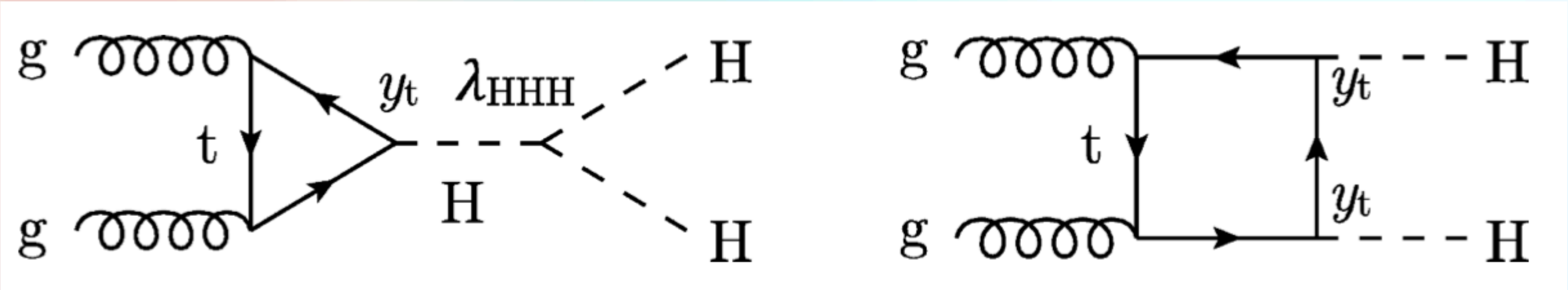
- In this analysis the non-resonant production of HH pairs was presented.

Process

$$pp \rightarrow HH$$

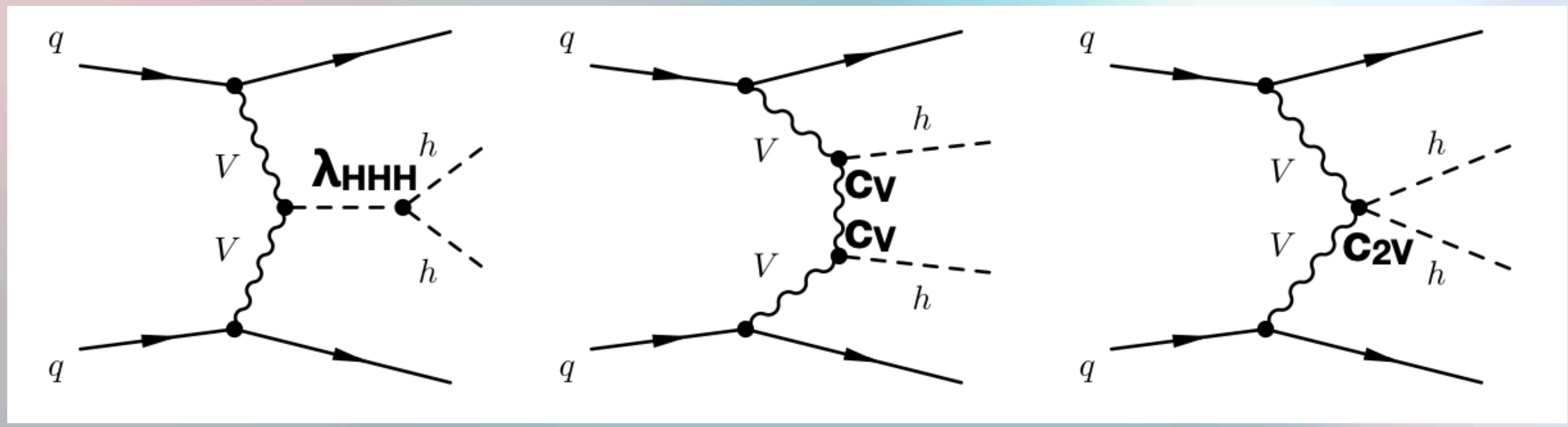
- Interest of the analysis:
  - Sensitive to trilinear self-coupling of H boson (λ_{HHH});
  - Sensitive to VVH coupling;
  - Sensitive to new particles contributing to contribute in virtual quantum loops

Dominant Production — gg Fusion (ggF) — $\sigma \approx 31\ fb$ @ 13 TeV



Distructive interference

Vector Boson Fusion (VBF) — $\sigma \approx 1.7\ fb$ @ 13 TeV



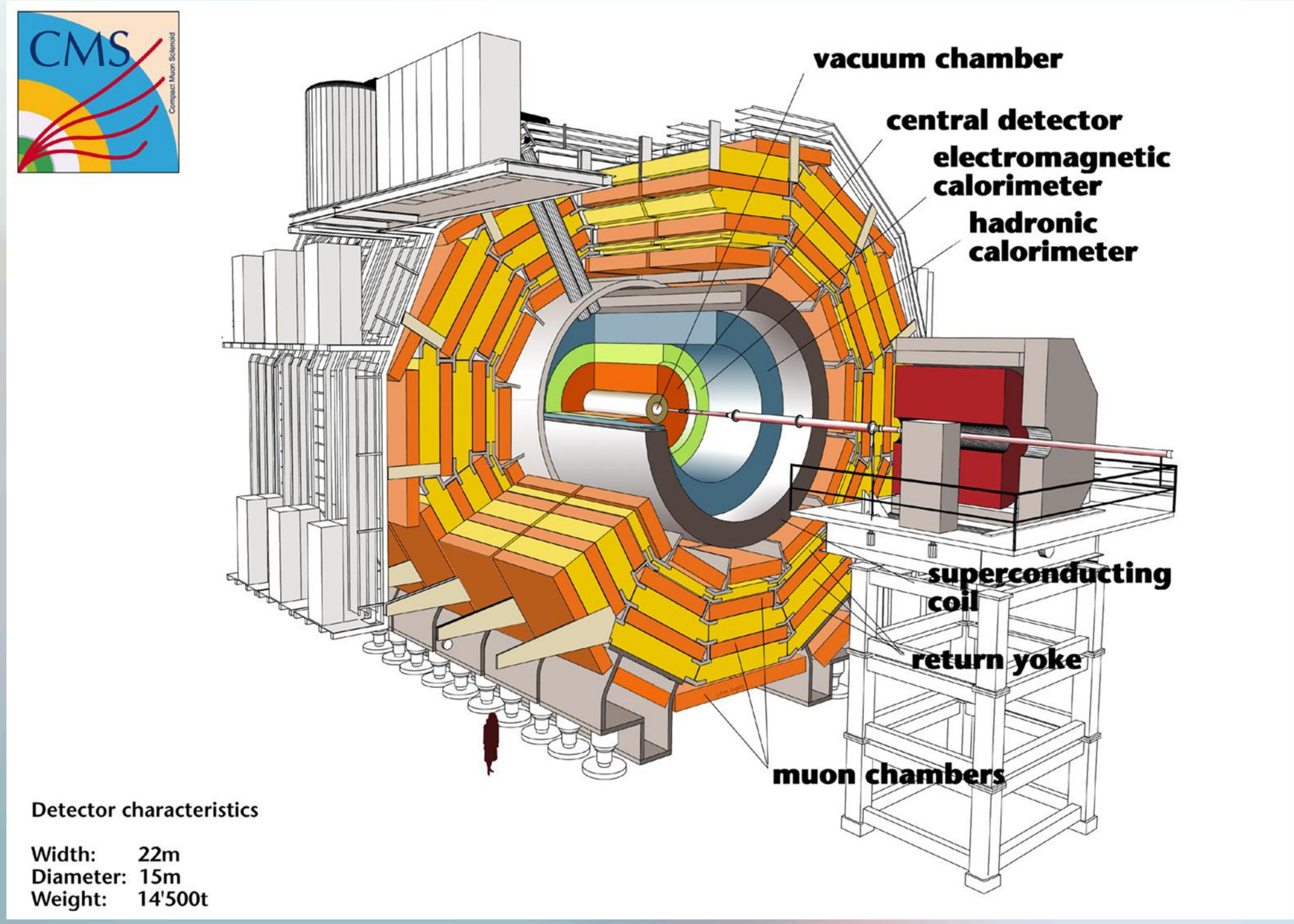
The process

A glance to the experimental setup

- The HH pair is not stable.
- Decay channel studied: $HH \to b\bar{b}\tau^+\tau^-$
  - 7.3% of branching fraction;
  - Precise Tau ID thanks to DEEPTau tool.
- $\tau^+\tau^-$ pair is not stable — analysis focused on the following decay modes:
  - $\tau_e\tau_h$ $(\tau_e = e\nu_e\nu_\tau)$
  - $\tau_\mu\tau_h$ $(\tau_\mu = \mu\nu_\mu\nu_\tau)$
  - $\tau_h\tau_h$
- HH decay products detected by CMS central system.

Process

$$pp \to HH$$



Data and simulated samples

Data taking and samples generation

- The analysis considers 2016+2017+2018 data taking periods;
- The total luminosity considered is 138 fb^{-1}.
- **Signal samples**: ✔
  - ggF simulated at LO with Madgraph and NLO with Poweg;
  - VBF simulated LO with Madgraph
    - Hadronisation simulated with Pythia;
    - Reconstruction: GEANT;
- **Background samples**: ✘
  - Main backgrounds: QCD (data driven), DY, $t\bar{t}$;
  - Other backgrounds considered: WJets + VV + VVV + singleH + single-top + ttV + ttVV
  - Standard DAS samples used;

Event reconstruction and selection

Triggers, objects definition and selection cuts

- Selection cuts are mostly common for all years.
- 2017-2018 selection cuts are strictly the same. If is present a difference regarding 2016, it is shown in parentheses.
- **Triggers**:
  - Each decay channel needs a different trigger;
    - *Lepton + τ_h channels*: single lepton + cross triggers (asks for a lepton + τ_h);
    - $\tau_h\tau_h$ channel: DiTau trigger;
  - Offline Pt cut higher than trigger threshold (trigger efficiency);

Online p_T trigger thresholds	single-e: $p_T > 25$ (32) GeV cross-e: electron $p_T > 24$ GeV, τ_h $p_T > 30$ GeV single-μ: $p_T > 22$ (24) GeV cross-μ: muon $p_T > 19$ (20) GeV, τ_h $p_T > 20$ (27) GeV ditau: $p_T > 35$ GeV, ditau VBF: $p_T > 20$ GeV
Offline p_T thresholds	online threshold +1 GeV (electrons and muons) online threshold +5 GeV (τ_h candidates)
η thresholds	electrons and muons: $\|\eta\| < 2.1$ tau: $\|\eta\| < 2.1$ (2.3) for ditau and cross (single) triggers
Lepton ID and isolation	tight electron BDT ID + isolation tight muon ID and isolation
τ_h ID ($\tau_e\tau_h$, $\tau_\mu\tau_h$ channels)	medium DeepTauVsJet tight DeepTauVsMu very-loose DeepTauVsEle
τ_h ID ($\tau_h\tau_h$ channel)	medium DeepTauVsJet very-loose DeepTauVsMu very-very-loose DeepTauVsEle
Distance to PV	$\|d_{xy}\| < 0.045$ cm (electrons and muons only) $\|d_z\| < 0.2$ cm
Pair selections	opposite-sign, $\Delta R > 0.5$

Event reconstruction and selection

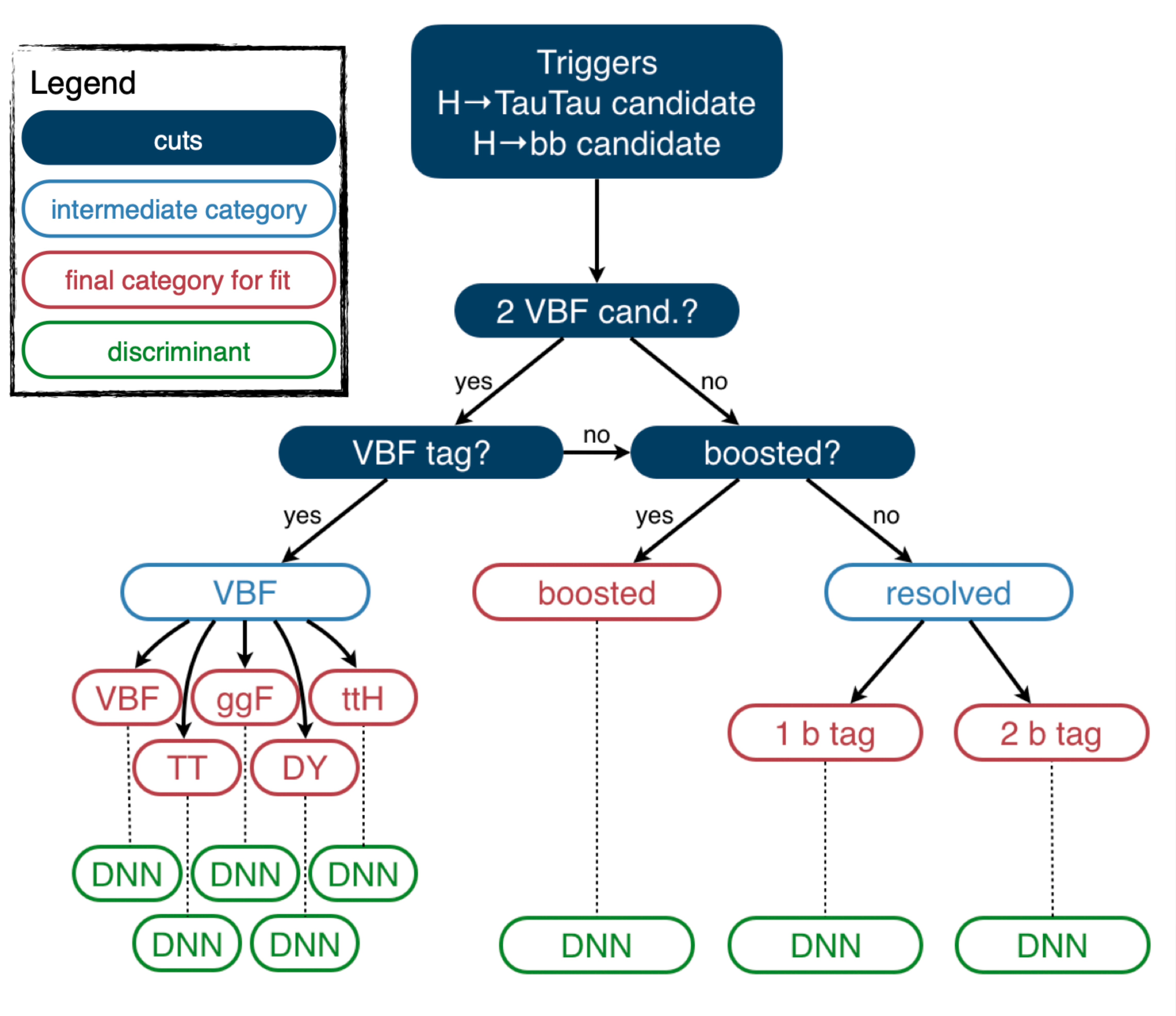
Triggers, objects definition and selection cuts

- Selection cuts are mostly common for all years.
- 2017-2018 selection cuts are strictly the same. If is present a difference regarding 2016, it is shown in parentheses.
- **Identification**:
  - *Leptons*:
    - Tight electron and Tight muons;
    - Isolation;
  - *Tau*: DEEPTAU (WP in the table)
- **Other relevant requirements**:
  - Geometrical acceptance ($|\eta|$);
  - Opposite sign leptons;
  - Distance to Primary Vertex.

Online p_T trigger thresholds	single-e: $p_T > 25$ (32) GeV cross-e: electron $p_T > 24$ GeV, τ_h $p_T > 30$ GeV single-μ: $p_T > 22$ (24) GeV cross-μ: muon $p_T > 19$ (20) GeV, τ_h $p_T > 20$ (27) GeV ditau: $p_T > 35$ GeV, ditau VBF: $p_T > 20$ GeV
Offline p_T thresholds	online threshold +1 GeV (electrons and muons) online threshold +5 GeV (τ_h candidates)
η thresholds	electrons and muons: $\|\eta\| < 2.1$ tau: $\|\eta\| < 2.1$ (2.3) for ditau and cross (single) triggers
Lepton ID and isolation	tight electron BDT ID + isolation tight muon ID and isolation
τ_h ID ($\tau_e\tau_h$, $\tau_\mu\tau_h$ channels)	medium DEEPTAUVSJET tight DEEPTAUVSMU very-loose DEEPTAUVSELE
τ_h ID ($\tau_h\tau_h$ channel)	medium DEEPTAUVSJET very-loose DEEPTAUVSMU very-very-loose DEEPTAUVSELE
Distance to PV	$\|d_{xy}\| < 0.045$ cm (electrons and muons only) $\|d_z\| < 0.2$ cm
Pair selections	opposite-sign, $\Delta R > 0.5$

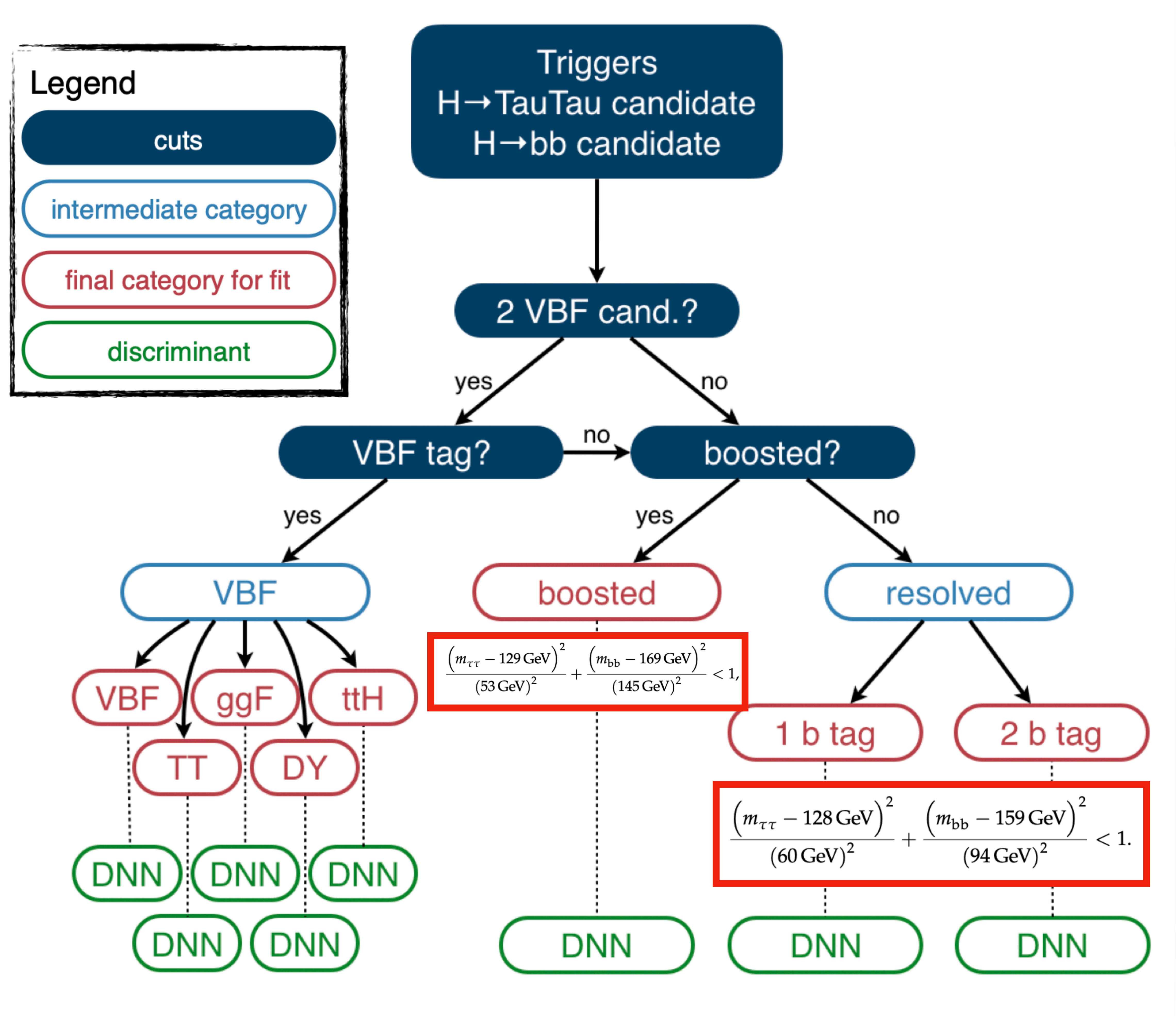
Event categorization

- Before passing to the multivariate part of the analysis, events are divided into 8 orthogonal categories.
- Categories are defined according to several criteria:
  - **VBF tag**: lets call $j_{1,2}$ the jets candidates.
    - Condition: $m_{j_1 j_2} > 500\ GeV,\ \Delta\eta > 3$
  - **VBF events are dived in sub-categories**:
    - Signal (VBF or ggF);
    - Backgrounds (TT, DY, ttH);
  - **No VBF events** are divided in 2 sub-categories:
    - **Resolved**: ΔR(bb)>0.4, reconstructed by AK4
    - **Boosted**: ΔR(bb)<0.4, reconstructed by AK8



Event categorization

- Before passing to the multivariate part of the analysis, events are divided into 8 orthogonal categories.
- Categories are defined according to several criteria:
  - **VBF tag**: lets call $j_{1,2}$ the jets candidates.
    - Condition: $m_{j_1 j_2} > 500\ GeV,\ \Delta\eta > 3$
  - **VBF events are dived in sub-categories**:
    - Signal (VBF or ggF);
    - Backgrounds (TT, DY, ttH);
  - **No VBF events** are divided in 2 sub-categories:
    - **Resolved**: ΔR(bb)>0.4, reconstructed by AK4
    - **Boosted**: ΔR(bb)<0.4, reconstructed by AK8



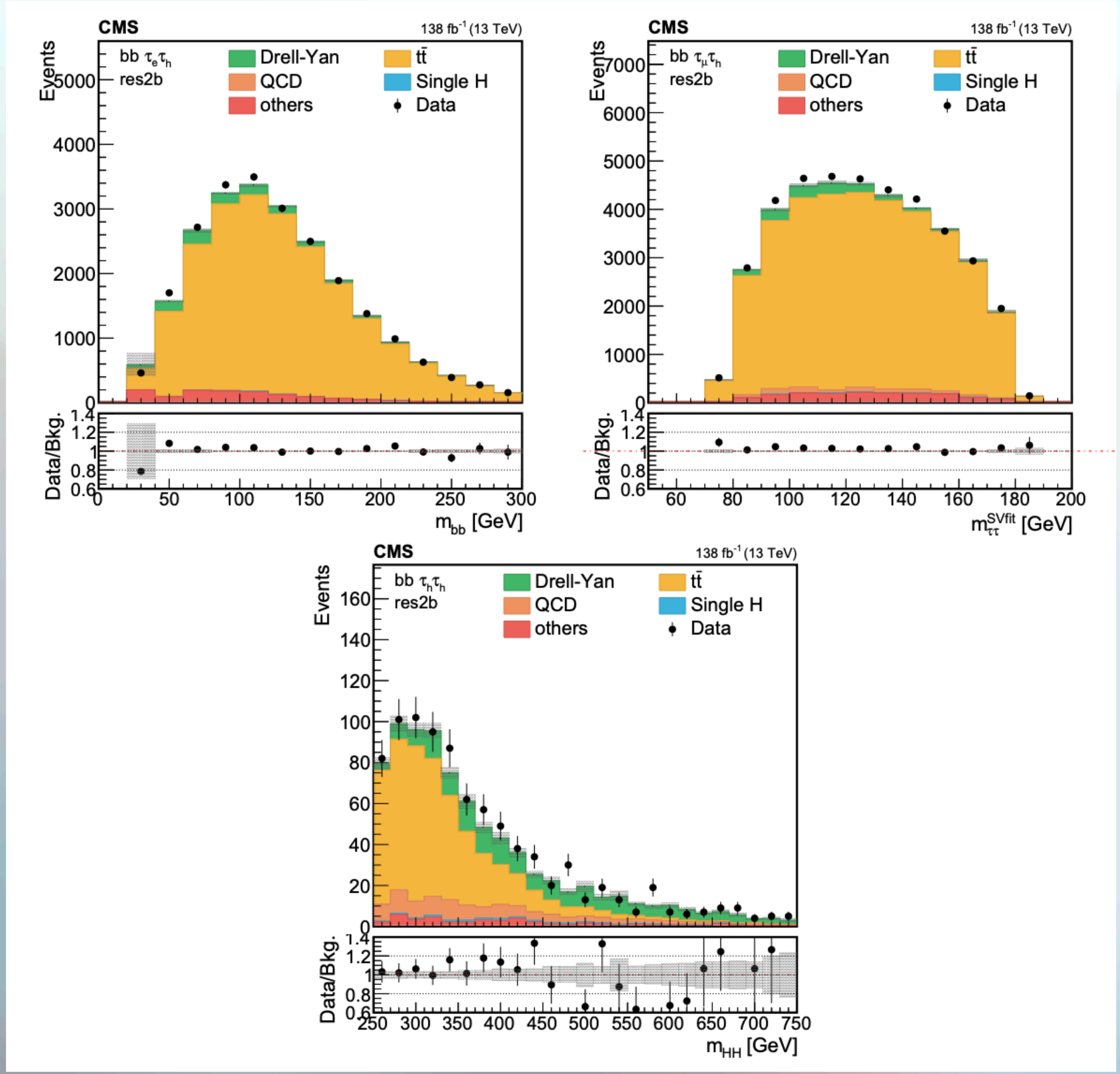
Multivariate analysis

- To best separate signal from background an MVA tool was developed, based on a DNN;
- The whole analysis is based on a single training (signal vs. backgrounds) with events from all years, channels and categories;
- Each event is subsequently associated to a single predictor;
  - Predictor closer to zero: BKG;
  - Predictor closer to one: SIGNAL;
  - The final distributions used in the signal extraction fit are obtained by inferring predictions of the trained network separately in each of the eight orthogonal categories (splitted by channel and year).

Multivariate analysis

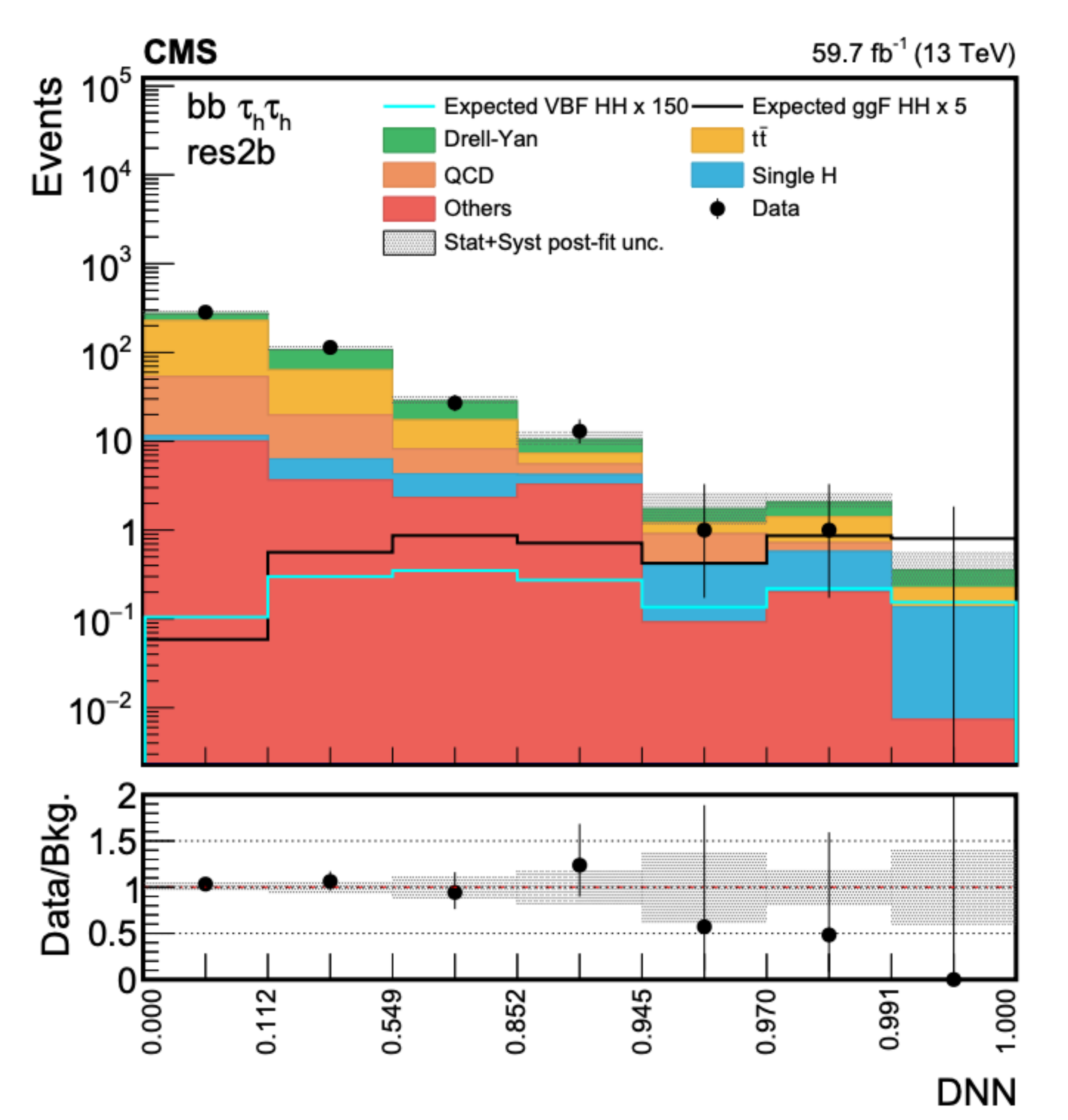
Input distributions

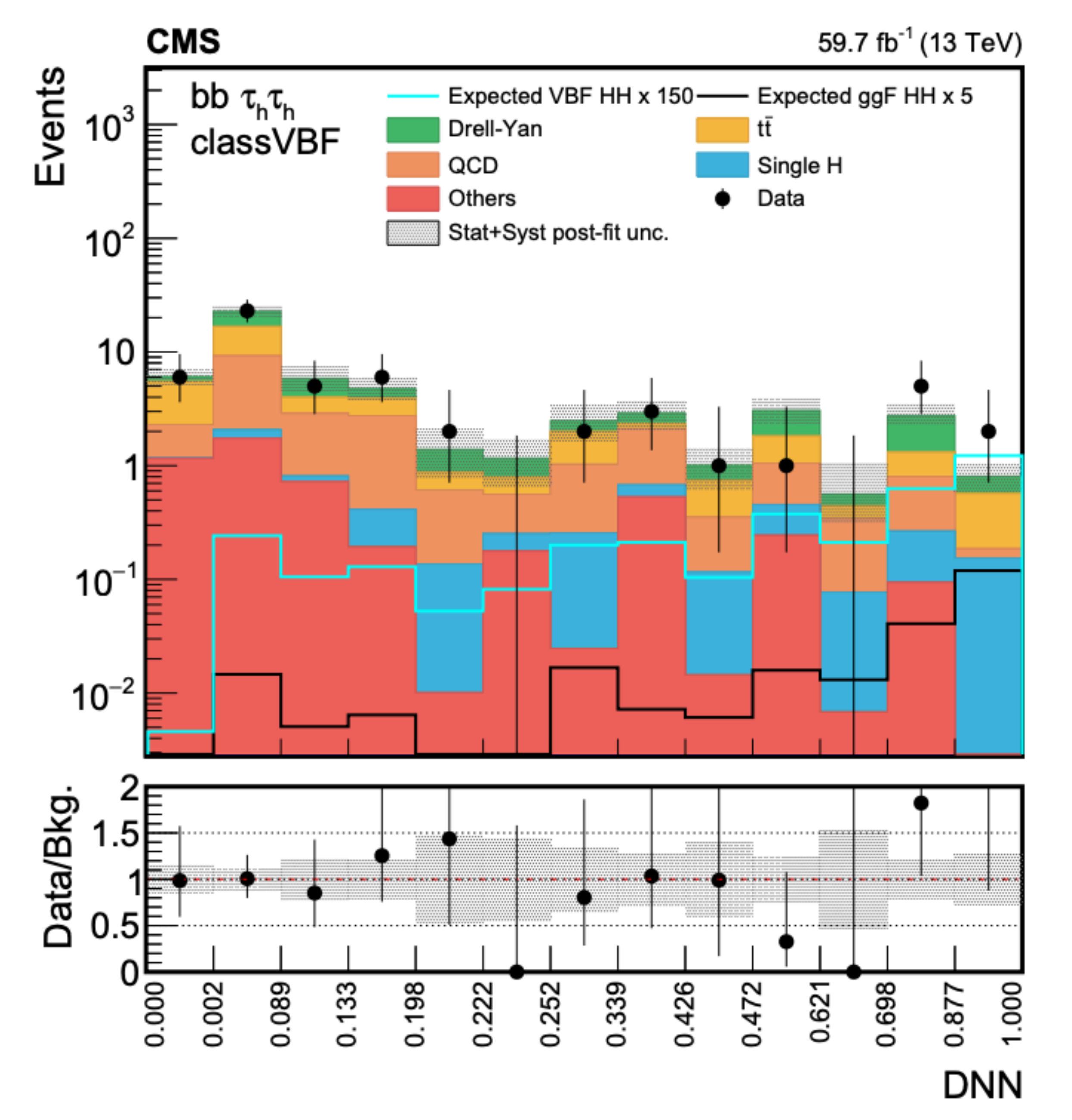
- The MVA tool was fed with 26 input distributions;
- The most discriminating distributions are shown here:
  - Invariant mass of the bb pair;
  - Invariant mass of the TauTau pair;
  - Invariant mass of the HH pair;
- Please note the good agreement between data and MC.



DNN output

- It was performed a likelihood fit of the DNN prediction to the data:
    - The fit takes into account 72 distributions (8 categories x 3 channels x 3 years);
- As an example, I show the DNN output for $\tau_h\tau_h$ channel (2018), considering the two most discriminant categories (VBF && ggF).
- Subsequently, limits @95% of CL are derived.



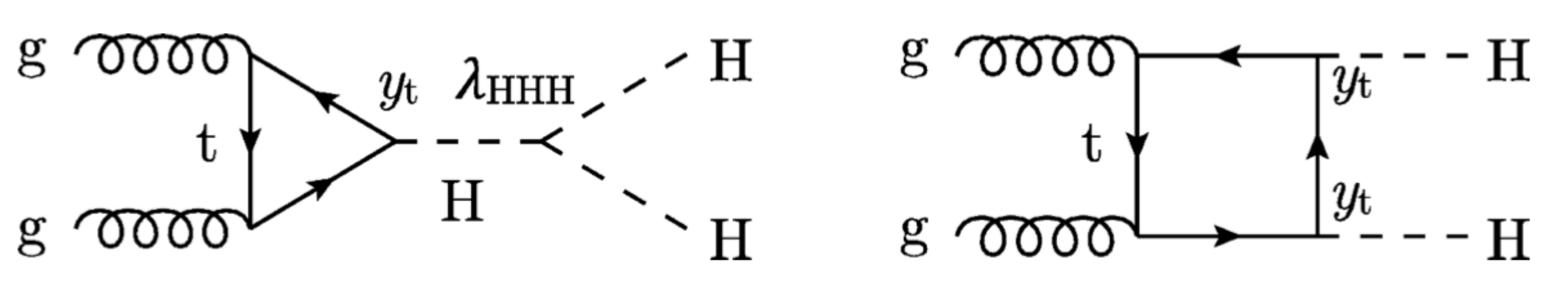


Results

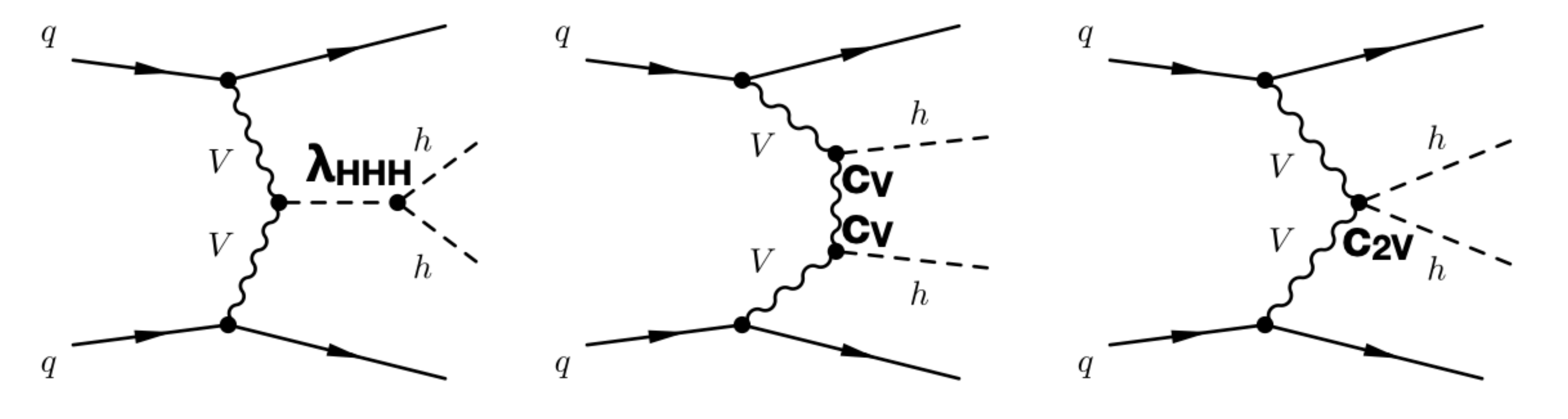
Limits on the SM cross section.

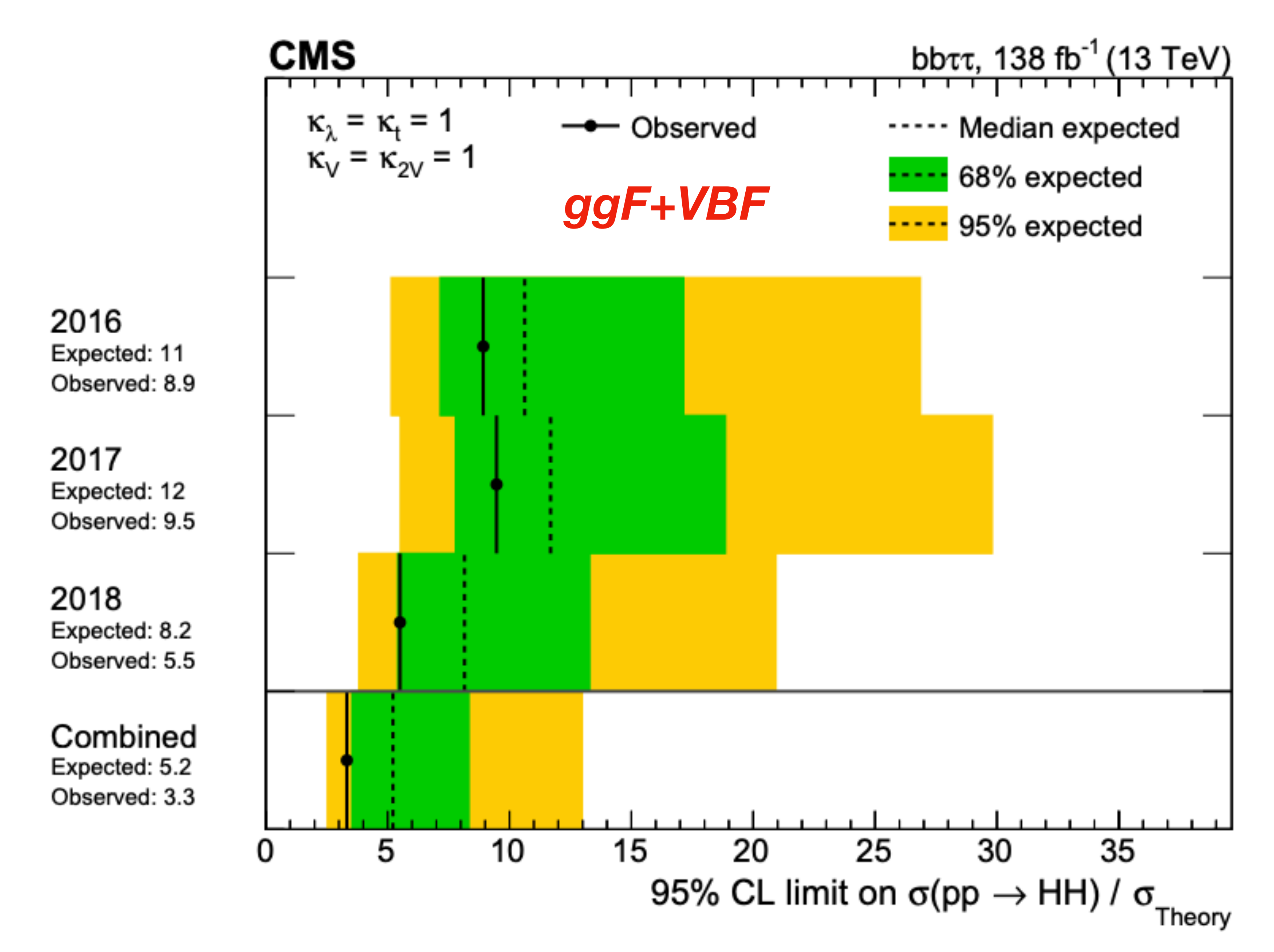
- Limits are derived both ggF+VBF HH production and VFB only production;
- **Signal strength observed**:
  - r = 3.3 (ggF + VBF)
  - r = 124 (VBF)

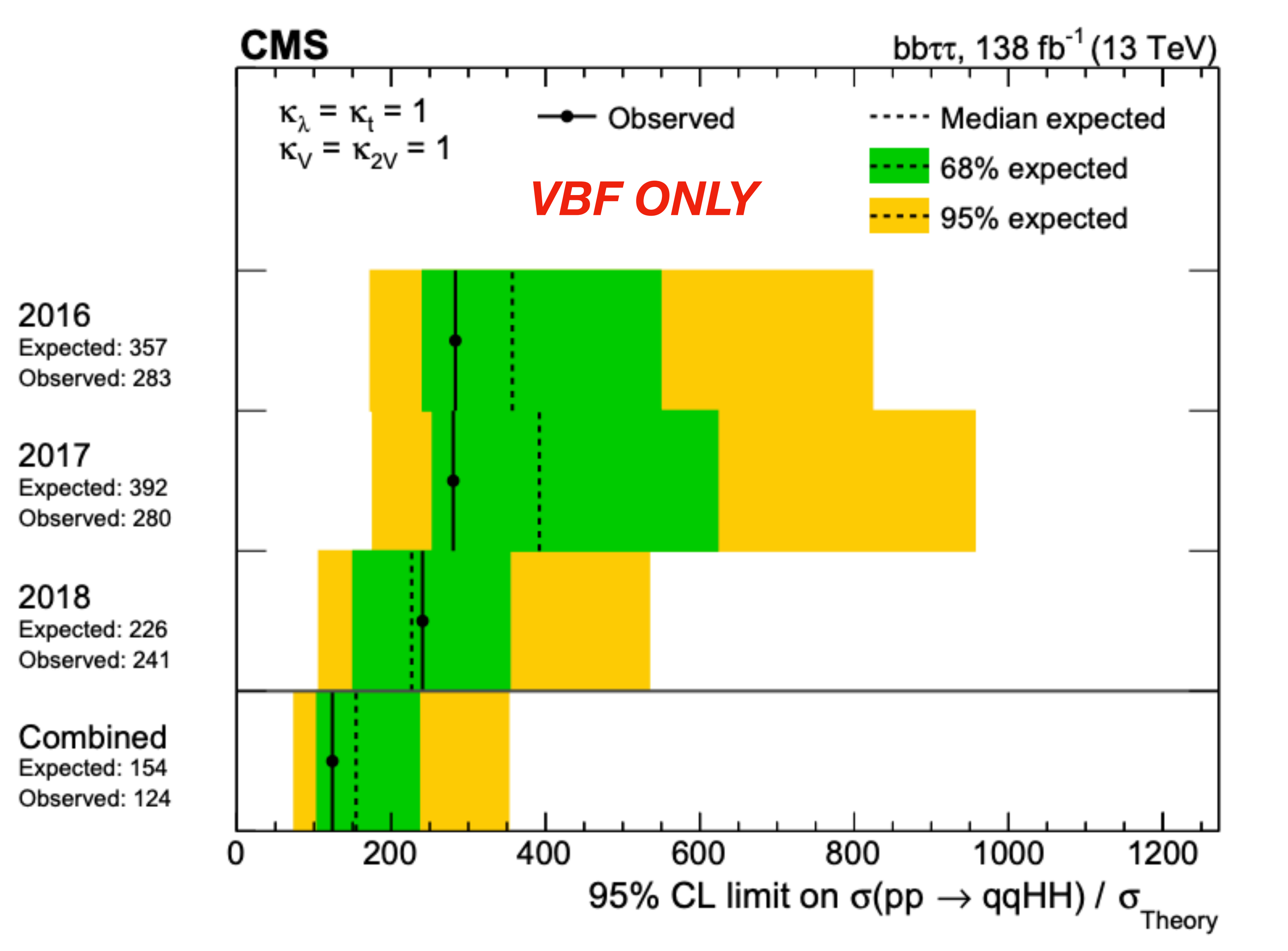
Dominant Production — gg Fusion (ggF) — $\sigma \approx 31\ fb$ @ 13 TeV



Vector Boson Fusion (VBF) — $\sigma \approx 1.7\ fb$ @ 13 TeV







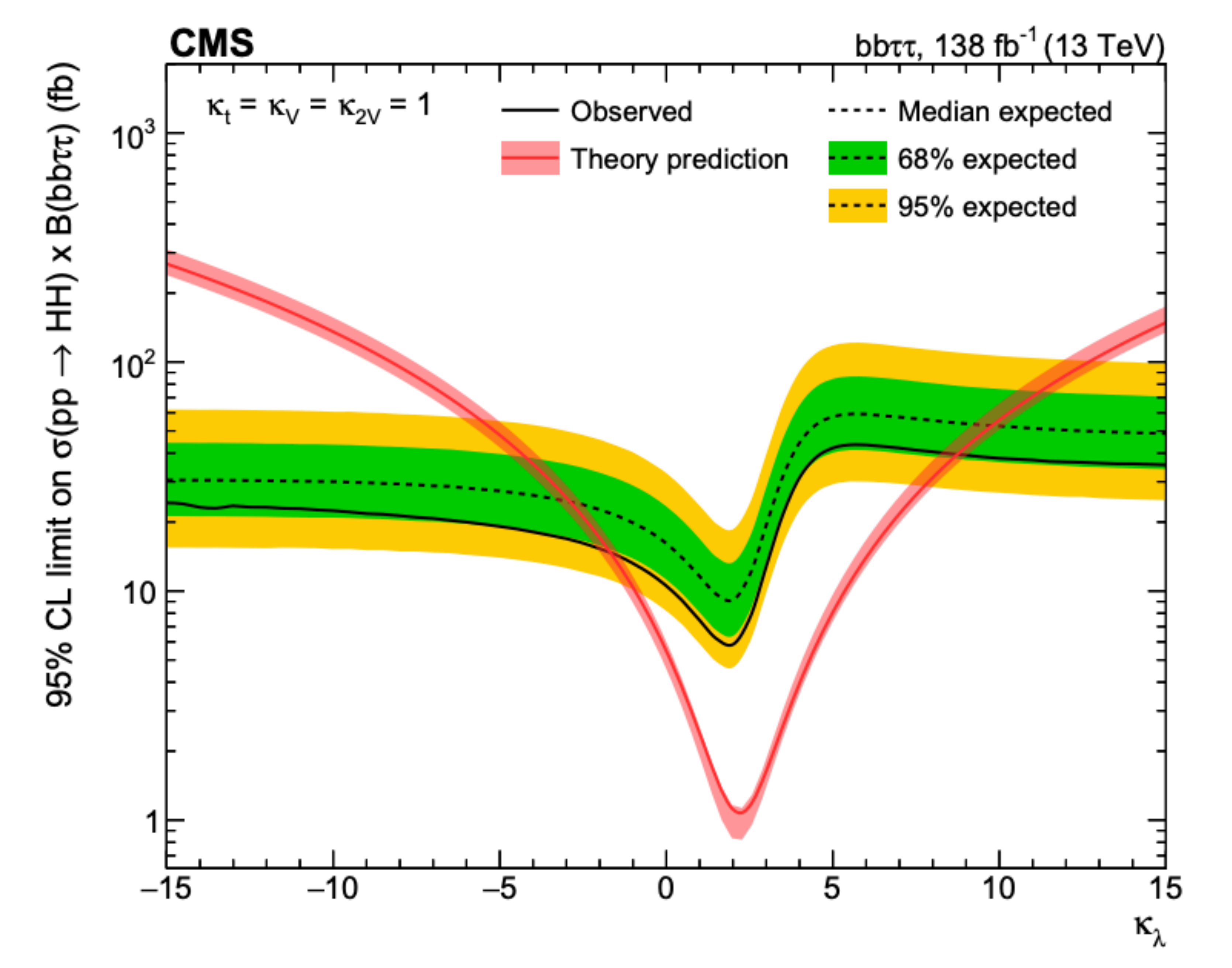
Results

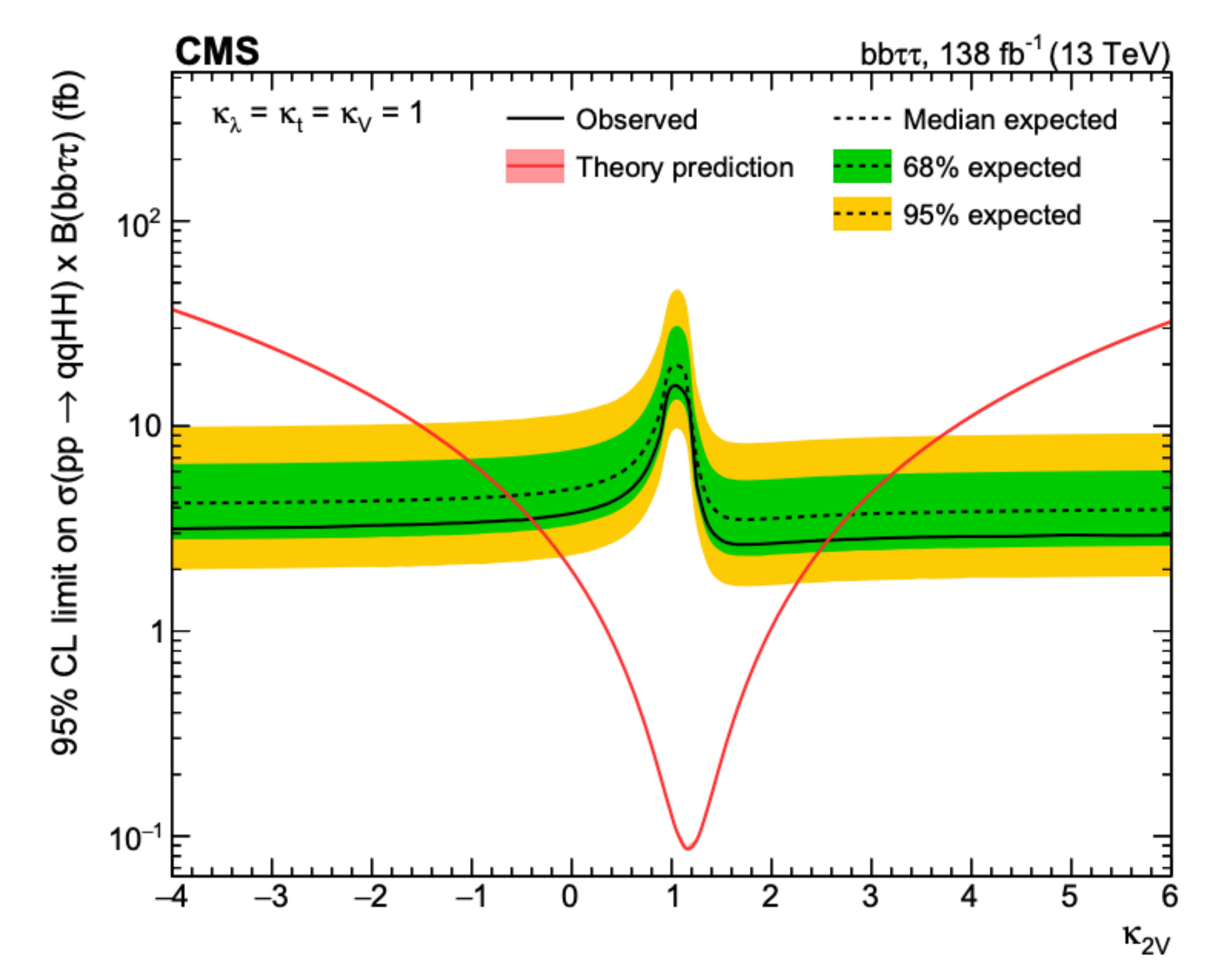
Trilinear HHH self coupling and VVH coupling.

- Fixing the value of $k_t,\ k_V,\ k_{2V}$ we can plot the dependency of σ as a function of k_λ (top plot)
- Comparing observed (expected) limit to the theory prediction one can derive the limits on λ_{HHH};
- Same reasoning can be applied to the bottom plot to derive c_{2V} limits.
- ***Observed limits***: -1.7 < k_λ < 8.7, -0.4 < k_{2V} < 2.6

Remind:

$$k_\lambda = \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}} \qquad k_{2V} = \frac{c_{2V}}{c_{2V}^{SM}}$$

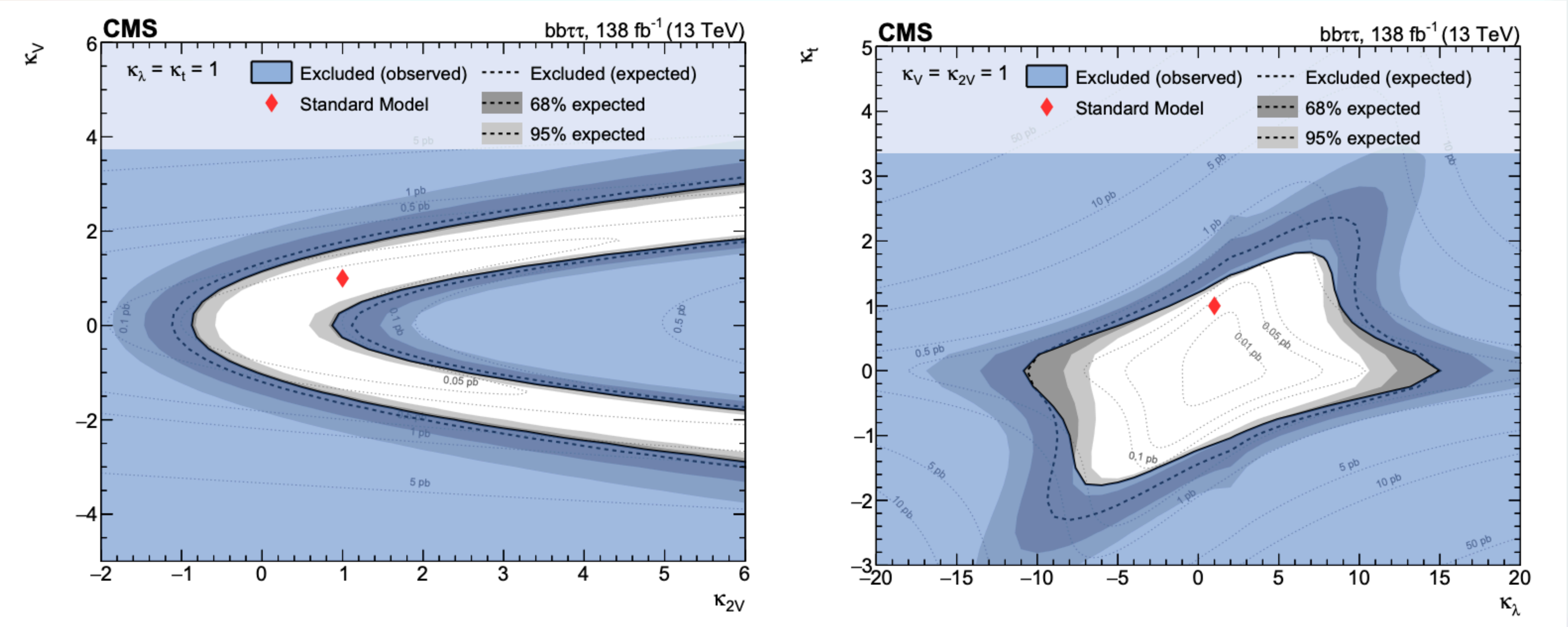




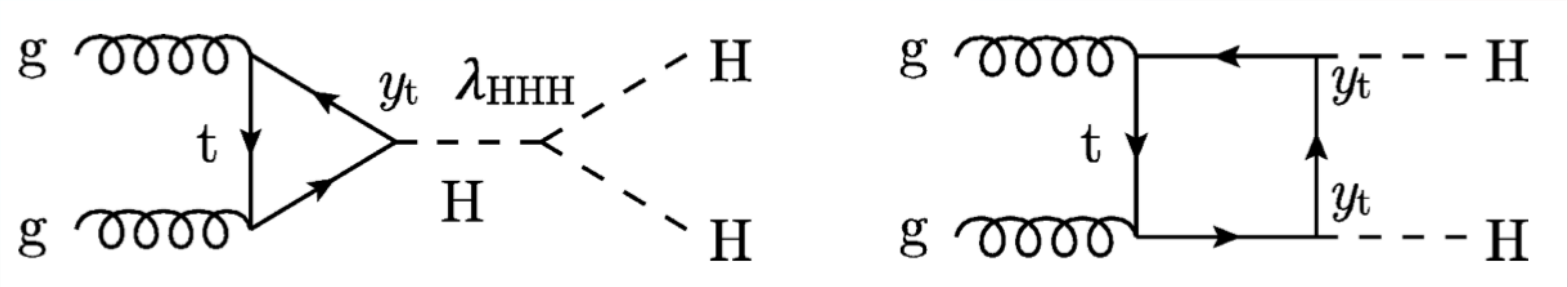
Results

Trilinear HHH self coupling and VVH coupling.

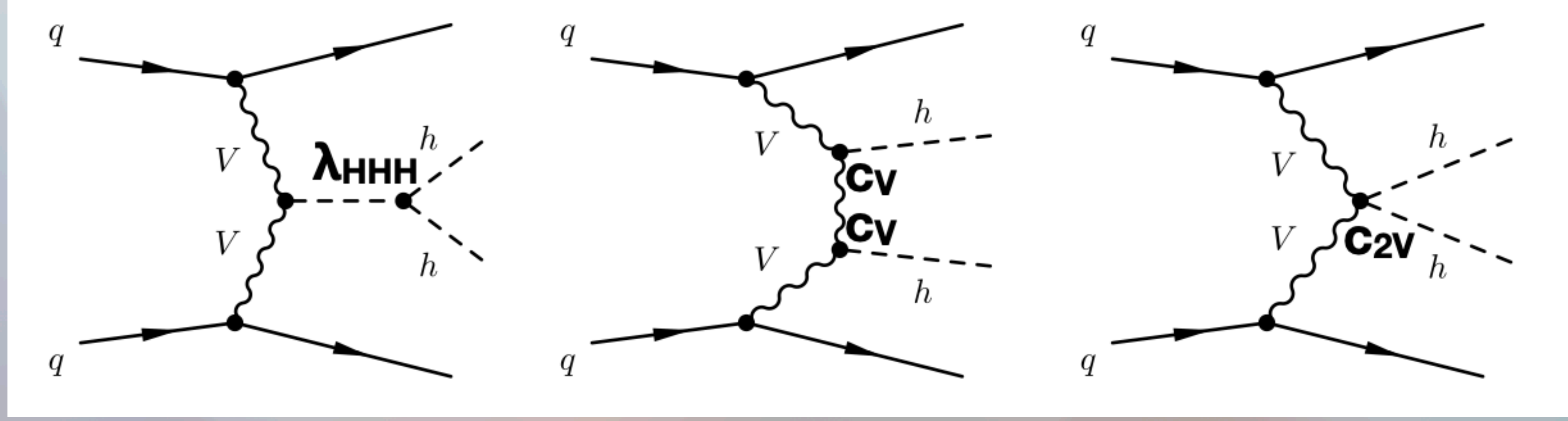
- At this point one can easily derive a bi-dimensional limit on $k_t\ vs.\ k_\lambda$ and $k_V\ vs.\ k_{2V}$.



Dominant Production — gg Fusion (ggF) — $\sigma \approx 31\ fb$ @ 13 TeV



Vector Boson Fusion (VBF) — $\sigma \approx 1.7\ fb$ @ 13 TeV

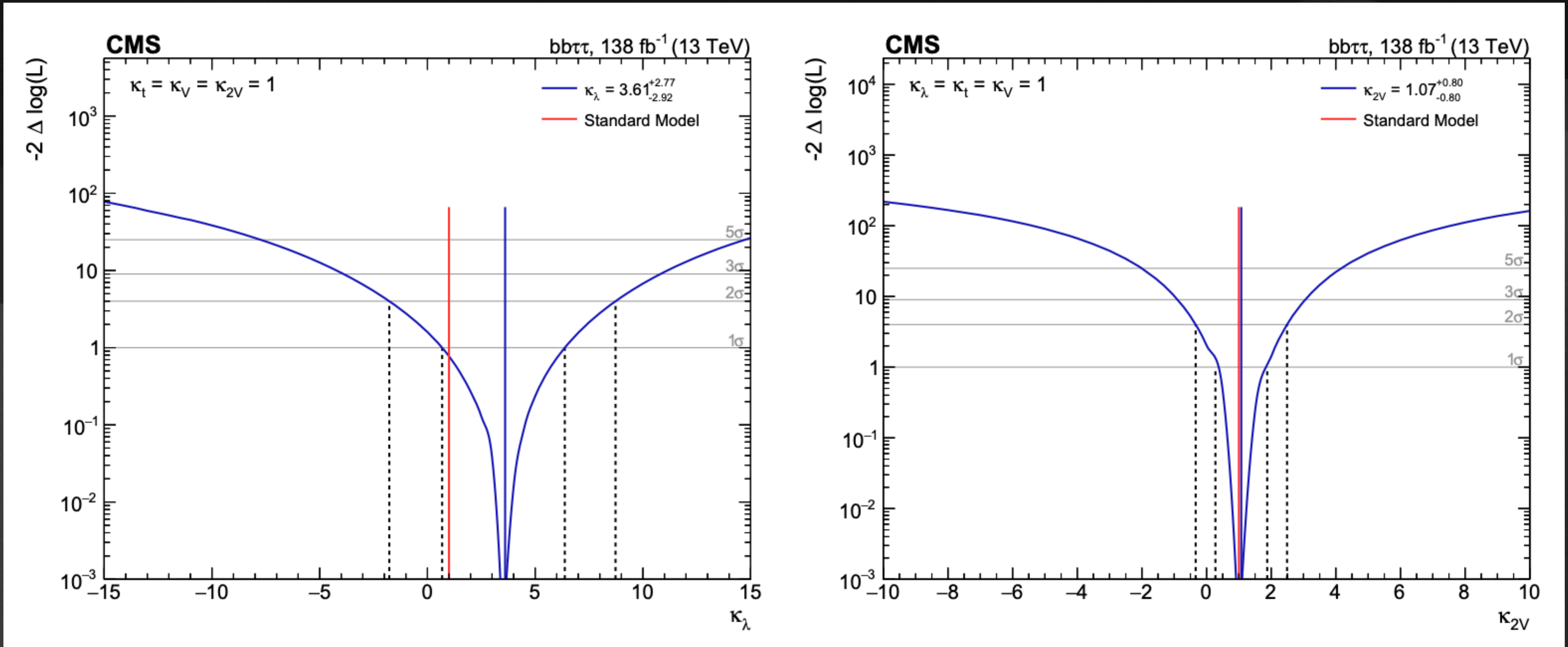


Remind:

$$k_\lambda = \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}} \qquad k_{2V} = \frac{c_{2V}}{c_{2V}^{SM}} \qquad k_V = \frac{c_V}{c_V^{SM}} \qquad k_t = \frac{y_t}{y_t^{SM}}$$

Thanks for the attention!

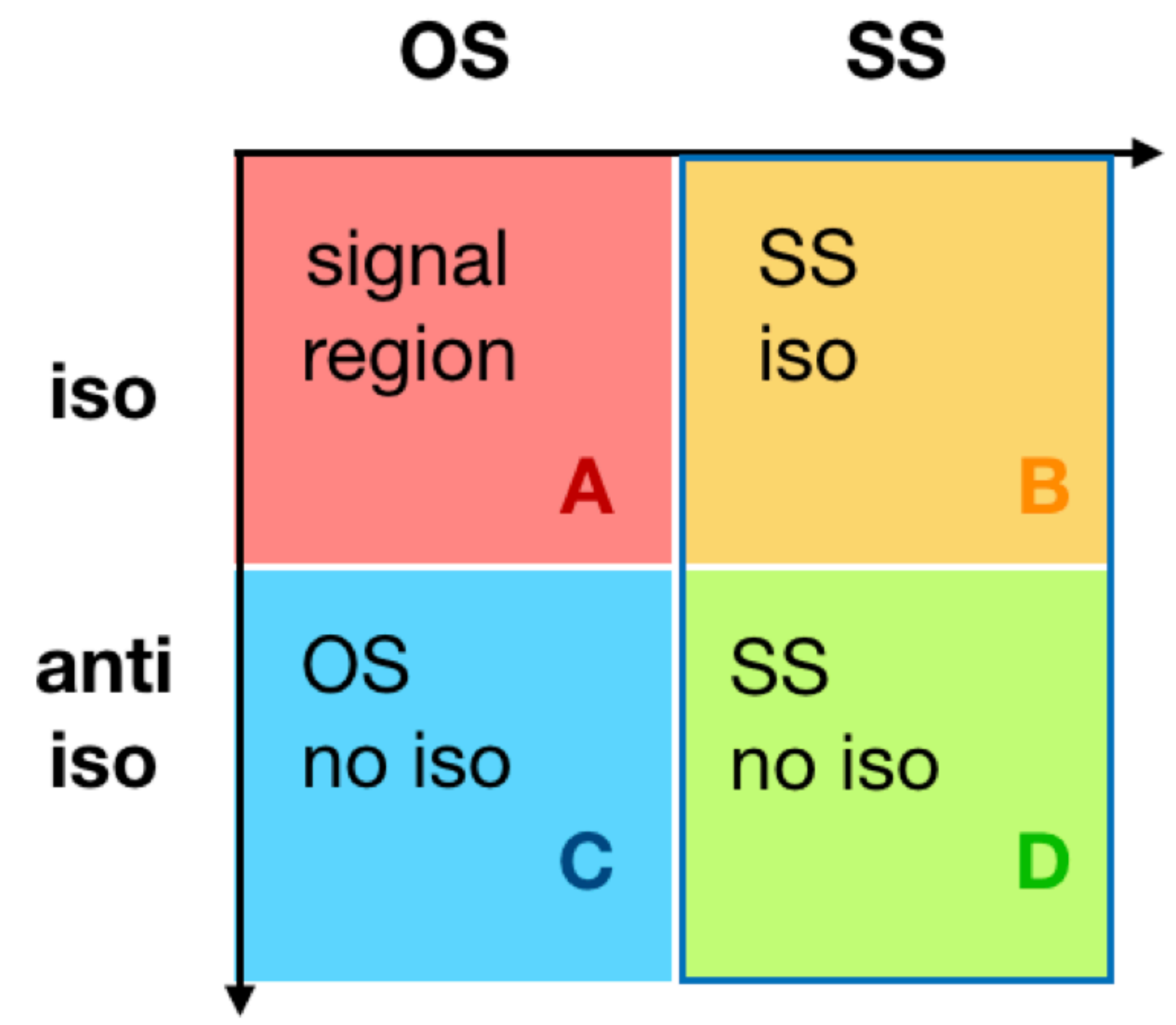
Best fit values for k_λ, k_{2V}



QCD data driven estimation

QCD:

Fully data-driven (ABCD method)



Signal region built from average of the two estimates:

$$\begin{aligned} A &= \boxed{C} * \boxed{B/D} \\ A &= \boxed{B} * \boxed{C/D} \end{aligned}$$

shape

correction to yield