北京大學
PEKING UNIVERSITY

Search for HH with bbμμ

Xiaohu Sun Botao Guo Licheng Zhang Zhe Li Yong Ban

Peking University

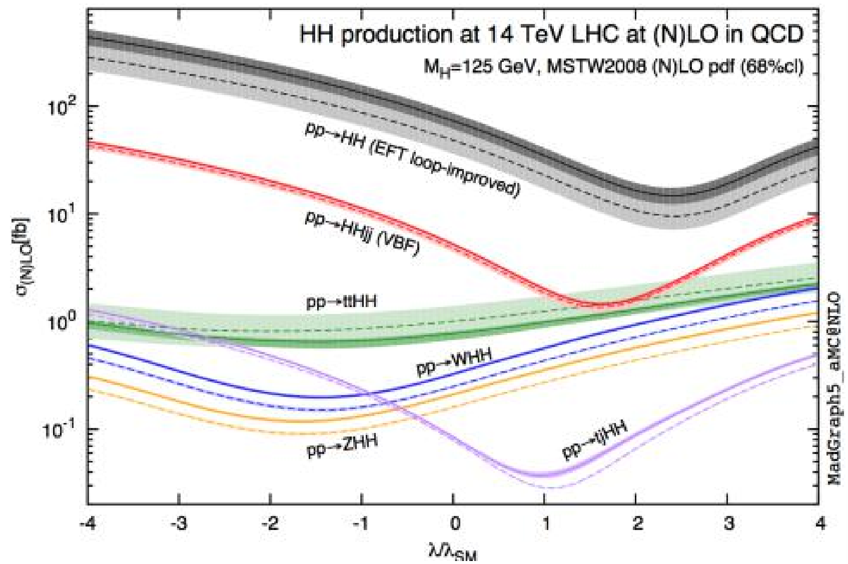
LHC Higgs WG 2022

28 Nov. 2022

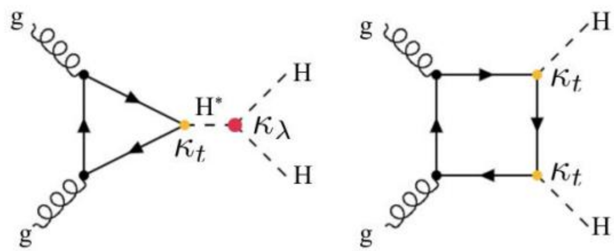
Based on arxiv: 2207.10912

Physics Motivation

- The trilinear self-coupling of the Higgs boson λ_{HHH} is only directly accessible via HH production at LHC
- The HH searches not fully explored the decays involving the second-generation fermions, such as HH $\rightarrow b b \mu\mu$

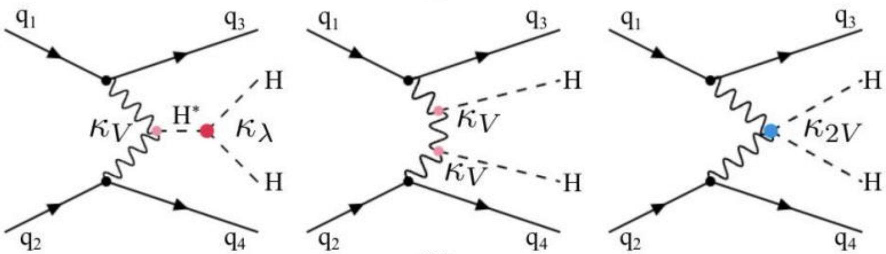


The dominant bkg of this process is DY + jets.
Other bkg, such as tt, single Higgs, which contains ggH, VBFH, ttH, ZH and bbH are also considered.



(a)

This rare process dominantly occurs via ggF. VBF is the second largest production mode. We only consider these two modes.

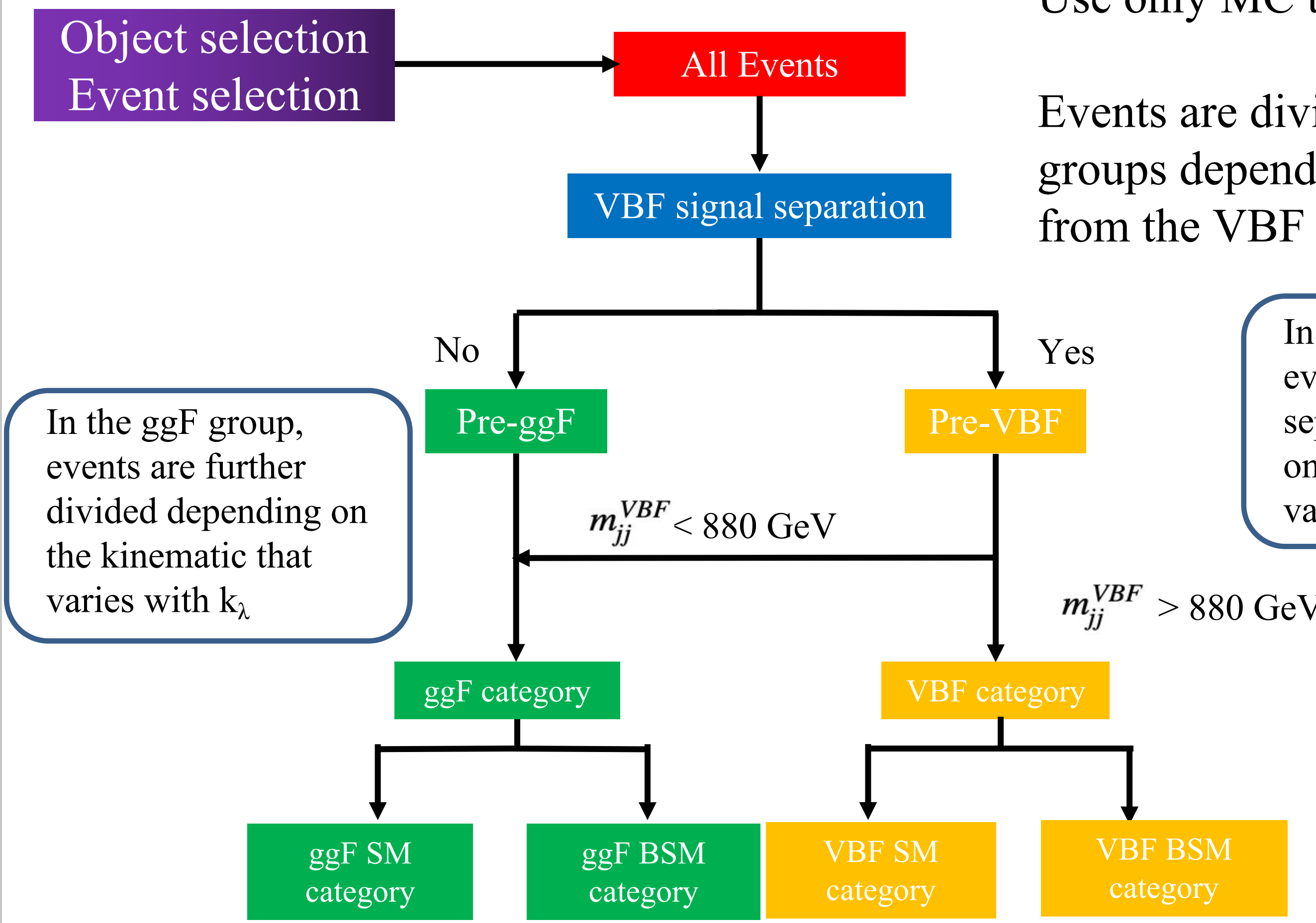


(b)

Analysis Strategy

Use only MC to do the search.

Events are divided into two groups depending on the jets from the VBF events topology.



In addition, in the two ggF categories, also used BDT approach (XGBoost) to suppress the bkg

In each signal category, do the bkg rejection to optimize the signal sensitivity vs bkg

Object selction and event selction

Object selction

For muon candidates:

- $p_T > 20$ GeV
- $|\eta| < 2.4$

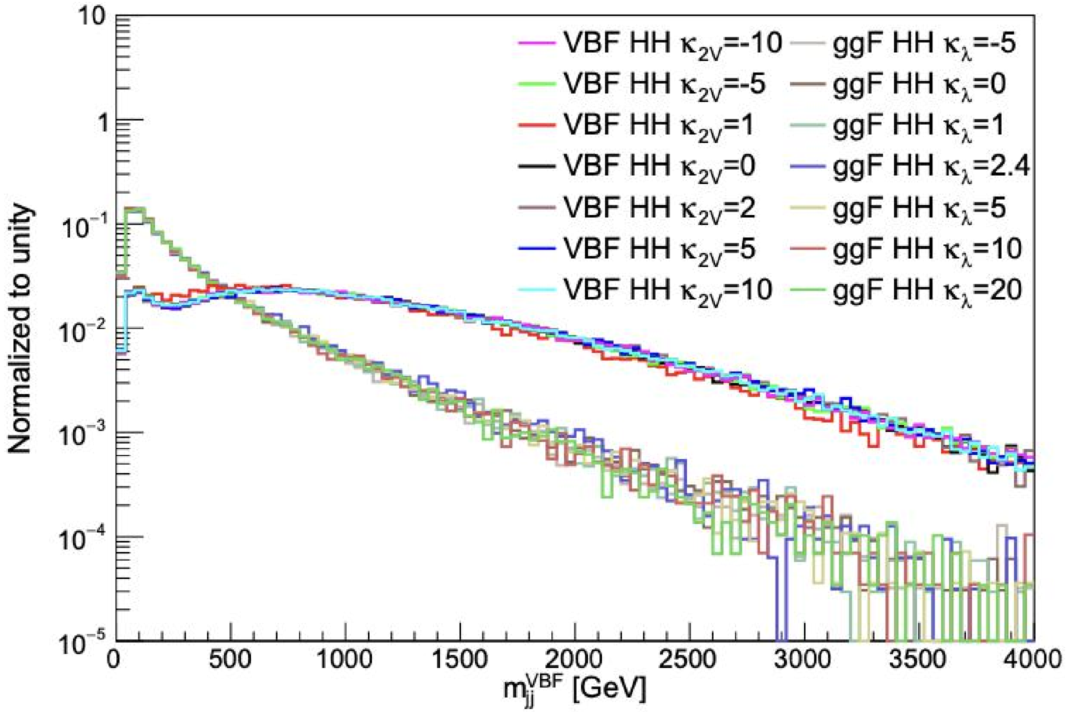
For jet candidates:

- $p_T > 20$ GeV
- $|\eta| < 4.7$ (2.4) for jet (b-tagged jet)

Event selction

- at least two oppositely charged muons
- 100 GeV $< m_{\mu\mu} <$ 180 GeV
- at least two bjets
- 70 GeV $< m_{bb} <$ 190 GeV

VBF and ggF categorization



To select events corresponding to HH production via VBF, additional requirements are imposed.

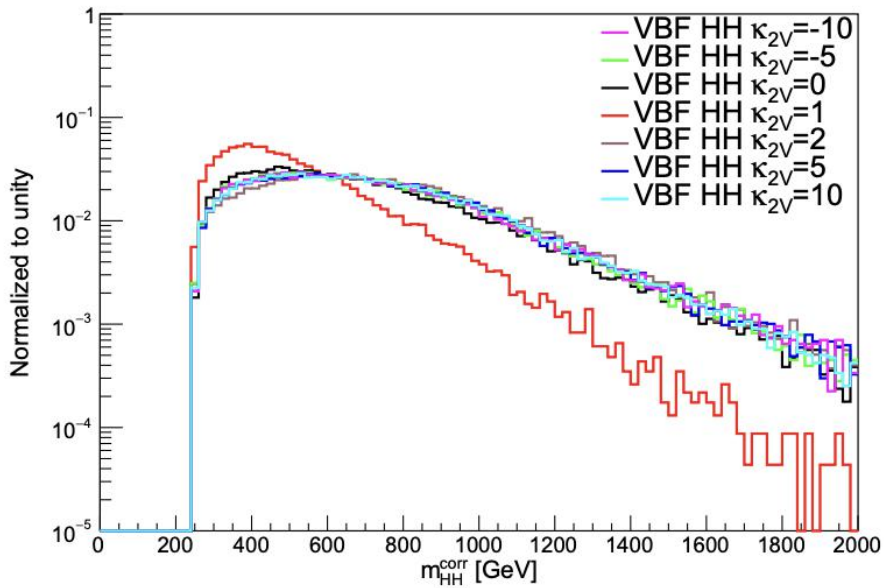
The VBF process is featured by the presence of two additional energetic jets.

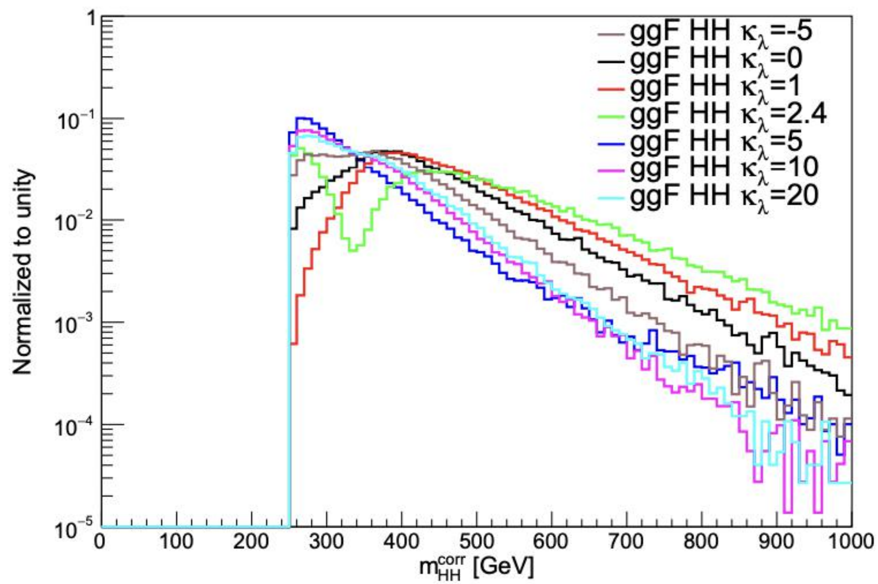
These VBF jets are expected to have a large pseudorapidity separation and a large dijet invariant mass.

The VBF events clearly have a much harder spectrum with respect to ggF. VBF as signal, ggF as bkg, use m_{jj}^{VBF} var to do optimization in order to maximize the separation of VBF and ggF signal (SM VBF and SM ggF as benckmark)

m_{jj}^{VBF} > 880 GeV as VBF category

Categorization in each VBF and ggF category





In the VBF group, events are then separated depending on the kinematic variation due to k_{2V}. The signals with anomalous k_{2V} values have a much harder spectrum than the SM VBF HH signal.

On the contrary, signal events generated with kλ close to 1 are more enriched in the high mass region. The other events enters more in the low mass region.

In VBF category

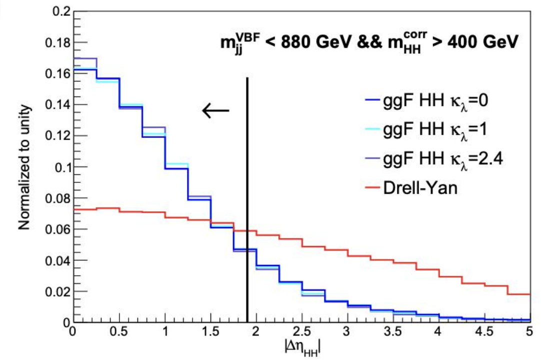
split the signal into 2 catrgories:
< 680 GeV : VBF SM category
> 680 GeV : VBF BSM category

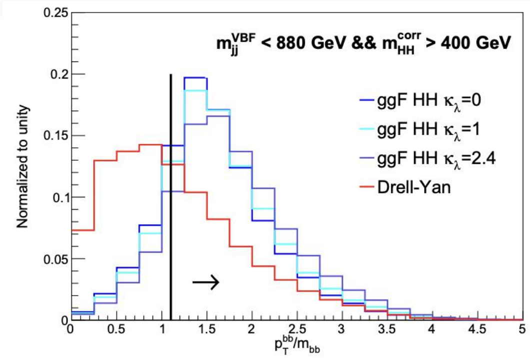
In ggF category

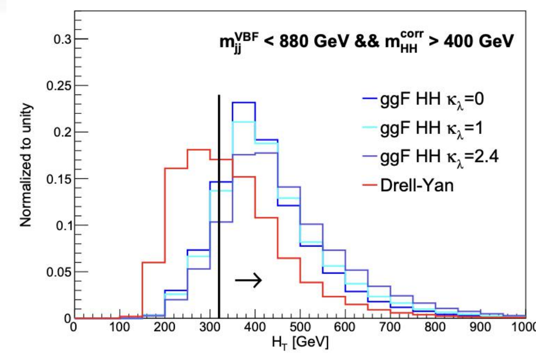
split the signal into 2 categories:
< 400 GeV : ggF BSM category
> 400 GeV : ggF SM category

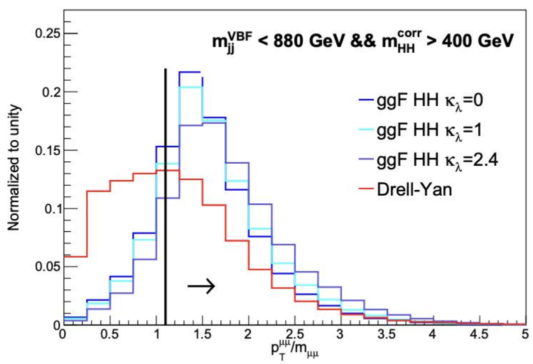
$$m_{\mathrm{HH}}^{\mathrm{corr}} = m_{\mathrm{HH}} - (m_{\mu\mu} - 125) - (m_{bb} - 125)$$

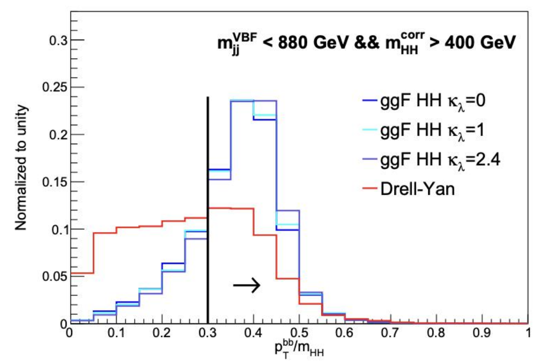
Bkg rejection in ggF SM category (Cutbased analysis)









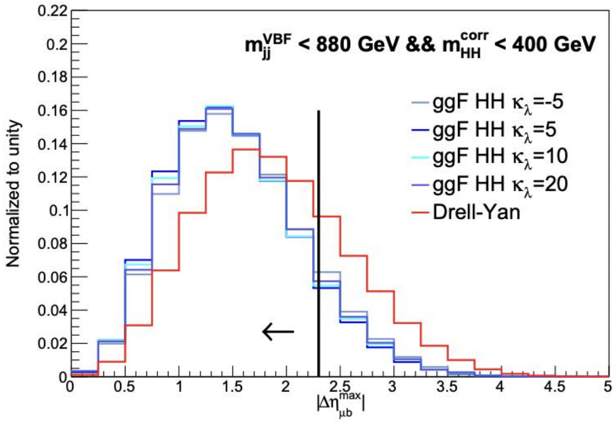


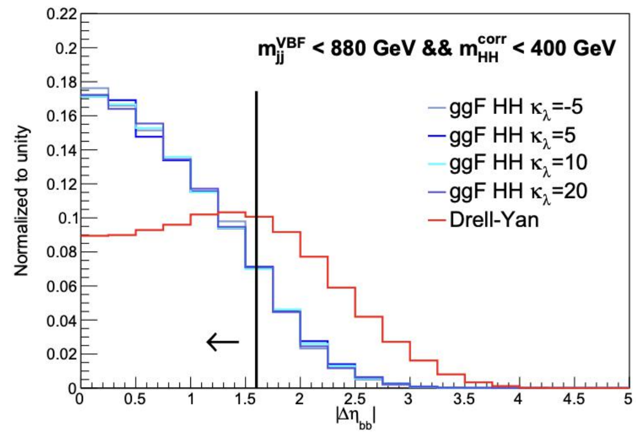
$$|\Delta\eta_{HH}| < 1.9$$
$$p_T^{bb}/m_{bb} > 1.1$$
$$p_T^{\mu\mu}/m_{\mu\mu} > 1.1$$
$$p_T^{bb}/m_{HH} > 0.3$$
$$H_T > 320 \text{ GeV}$$

Five discriminating variables are chosen. The |ΔηHH|, both the signal and the DY events are energetic and are mostly having the bb and μμ systems back-to-back. The signal events tend to be more transverse resulting in smaller values, while the DY events are less transverse leading to a larger spread in this variable, as shown in upper left plots.

The relative pT variables and HT, are very effective in separating the signal and the DY background events.

Bkg rejection in ggF BSM category (Cutbased analysis)

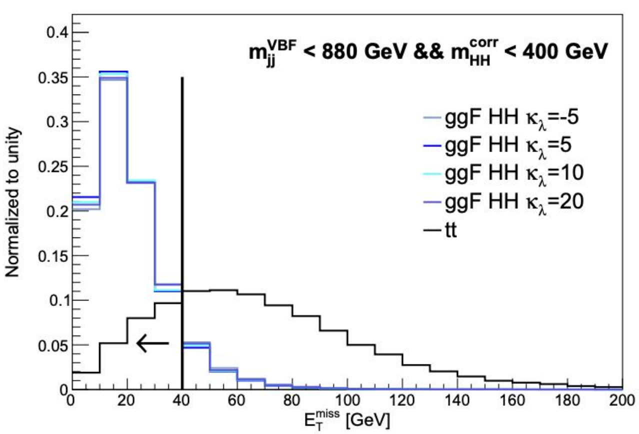




$$|\Delta\eta_{\mu b}^{max}| < 2.3$$

$$|\Delta\eta_{bb}| < 1.6$$

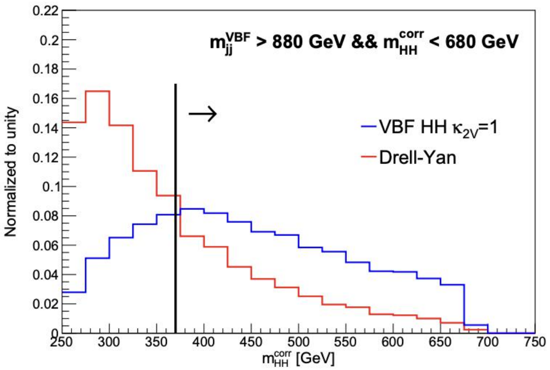
$$E_{\mathrm{T}}^{miss} < 40 \text{ GeV}$$

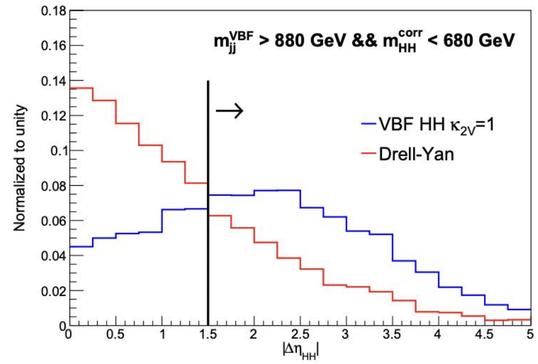


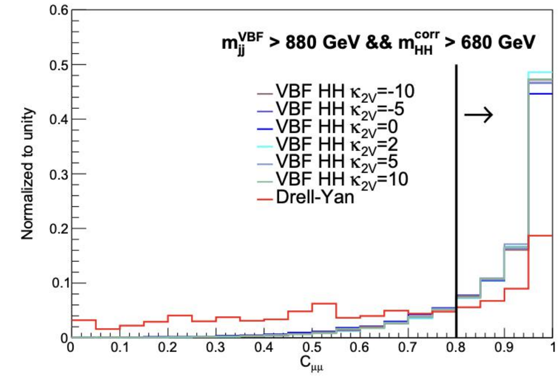
There are two outstanding variables that can improve the signal significance.

Different than other categories, there emerge a lot of tt events. Almost all of the top quark decay through the weak interaction, producing a W boson and a bottom quark. Since the W boson decays leave neutrino that almost does not interact with the detector.
E_T^{miss} is chosen to reject the tt events.

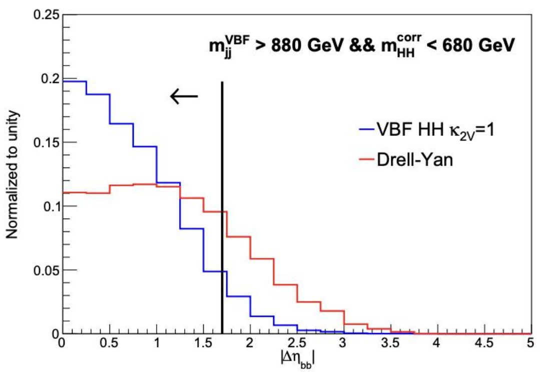
Bkg rejection in VBF SM, BSM category (Cutbased analysis)

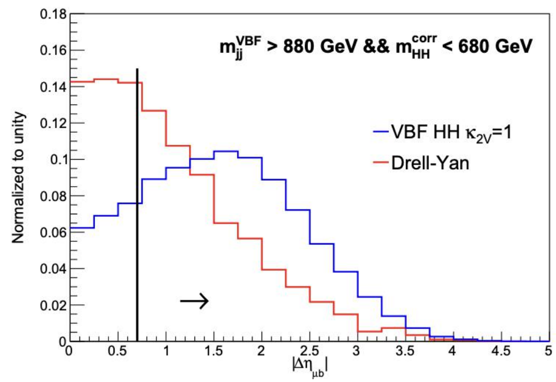






VBF BSM category





VBF SM category

$$m_{HH}^{corr} > 370 \text{ GeV}$$
$$|\Delta\eta_{HH}| > 1.5$$
$$|\Delta\eta_{bb}| < 1.7$$
$$|\Delta\eta_{\mu b}| > 0.7$$

$$C_{(H\to\mu\mu)} > 0.8$$

In the VBF SM category, there found 4 variables that improve the significance in a sizable way.

In the VBF BSM category, has the lowest statistics among all due to m_{HH}^{corr} > 680 GeV leave little room for optimization. Only one variable Cμμ is chosen.

$$C_{\mu\mu} = e^{[-\frac{4}{(\eta_1^{VBF}-\eta_2^{VBF})^2}(\eta^{\mu\mu}-\frac{\eta_1^{VBF}+\eta_2^{VBF}}{2})^2]}$$

Bkg rejection (Cutbased analysis)

Category	Variable Cut	ϵ_{Signal}	$\epsilon_{\text{DY+tt}}$
ggF SM	$\lvert\Delta\eta_{\text{HH}}\rvert < 1.9$ $p_{\text{T}}^{bb}/m_{bb} > 1.1$ $p_{\text{T}}^{\mu\mu}/m_{\mu\mu} > 1.1$ $p_{\text{T}}^{bb}/m_{\text{HH}} > 0.3$ $H_{\text{T}} > 320$ GeV	53% ($\kappa_\lambda = 1$)	11%
ggF BSM	$\lvert\Delta\eta_{\mu b}^{max}\rvert < 2.3$ $\lvert\Delta\eta_{bb}\rvert < 1.6$ $E_{\text{T}}^{miss} < 40$ GeV	55% ($\kappa_\lambda = 5$)	13%
VBF SM	$m_{\text{HH}}^{\text{corr}} > 370$ GeV $\lvert\Delta\eta_{\text{HH}}\rvert > 1.5$ $\lvert\Delta\eta_{bb}\rvert < 1.7$ $\lvert\Delta\eta_{\mu b}\rvert > 0.7$	39% ($\kappa_{2\text{V}} = 1$)	3%
VBF BSM	$C_{\mu\mu} > 0.8$	69% ($\kappa_{2\text{V}} = 10$)	15%

Category	m_{jj}^{VBF}(GeV)	m_{HH}^{corr}(GeV)
ggF SM	< 880	> 400
ggF BSM	< 880	< 400
VBF SM	> 880	< 680
VBF BSM	> 880	> 680

List the cuts used in defining the above four categories.

Summary of the optimized cuts for background suppression and the corresponding efficiencies in all the four categories.

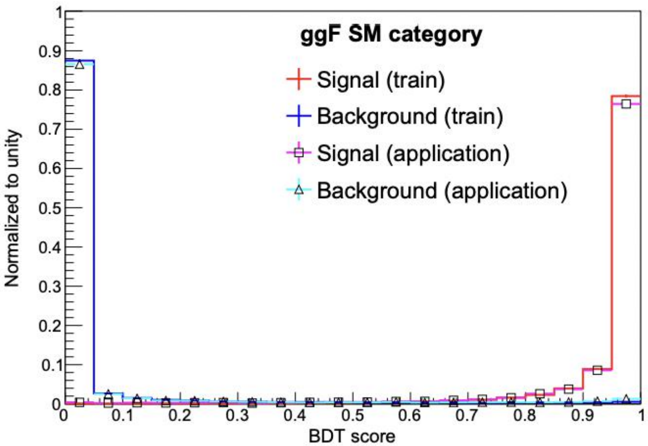
Bkg rejection (BDT analysis)

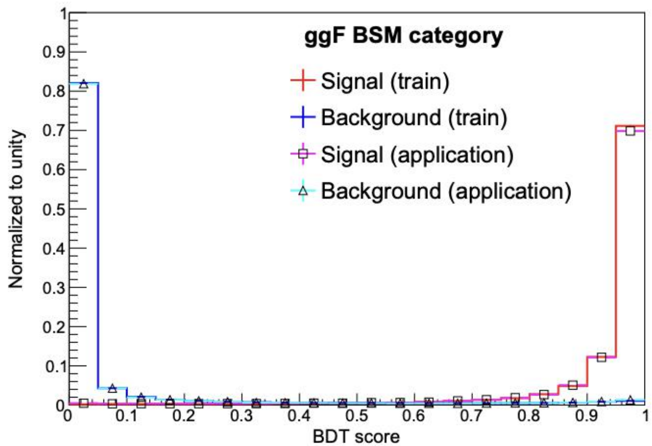
Input Variable	Category ggF SM	Category ggF BSM
$p_{\mathrm{T}}^{\mu1}$,$p_{\mathrm{T}}^{\mu2}$,p_{T}^{b1},p_{T}^{b2}	✓	✓
$E_{\mu1}$,$E_{\mu2}$,E_{b1},E_{b2}	✓	
$\eta^{\mu1}$,$\eta^{\mu2}$	✓	
η^{b1},η^{b2}	✓	✓
η_{j1}^{VBF}		✓
$E_{\mu\mu}$,E_{bb},$\eta_{\mu\mu}$,η_{bb},$\cos\theta_{\mu\mu}$,$\cos\theta_{bb}$	✓	
$p_{\mathrm{T}}^{\mu\mu}$,p_{T}^{bb},$m_{\mu\mu}$,m_{bb}	✓	✓
m_{HH},$m_{\mathrm{HH}}^{\mathrm{corr}}$	✓	
$p_{\mathrm{T}}^{b1}/m_{bb}$,$p_{\mathrm{T}}^{b2}/m_{bb}$,$p_{\mathrm{T}}^{bb}/m_{bb}$,$p_{\mathrm{T}}^{\mu1}/m_{\mu\mu}$,$p_{\mathrm{T}}^{\mu2}/m_{\mu\mu}$	✓	
$p_{\mathrm{T}}^{bb}/m_{bb}$,$p_{\mathrm{T}}^{bb}/m_{\mathrm{HH}}$,$p_{\mathrm{T}}^{\mu\mu}/m_{\mu\mu}$,$p_{\mathrm{T}}^{\mu\mu}/m_{\mathrm{HH}}$	✓	✓
H_{T},$p_{\mathrm{T}}^{\mathrm{HH}}$,$p_{\mathrm{T}}^{\mu\mu}/p_{\mathrm{T}}^{bb}$	✓	✓
E_{T}^{miss},η^{miss}	✓	✓
$\|\Delta\eta_{\mathrm{HH}}\|$,$\|\Delta\eta_{\mu b}\|$,$\|\Delta\eta_{\mu b}^{max}\|$,$\|\Delta\eta_{\mu b}^{other}\|$	✓	✓
$\|\Delta\eta_{bb}\|$,$\|\Delta\eta_{\mu\mu}\|$	✓	✓
$\|\Delta R_{\mathrm{HH}}\|$,$\|\Delta R_{\mu b}\|$,$\|\Delta R_{bb}\|$,$\|\Delta R_{\mu\mu}\|$	✓	✓
$\|\Delta R_{\mu b}^{min}\|$,$\|\Delta R_{\mu b}^{other}\|$,$\|\Delta R_{jj}^{\mathrm{VBF}}\|$	✓	
$\|\Delta\phi_{\mathrm{HH}}\|$,$\|\Delta\phi_{\mu b}\|$,$\|\Delta\phi_{bb}\|$,$\|\Delta\phi_{jj}^{\mathrm{VBF}}\|$	✓	✓
$\|\Delta\phi_{\mu\mu}\|$	✓	

In the ggF SM and ggF BSM category, we train a huge amount of shallow decision trees and form them into an output with a strong separation power.

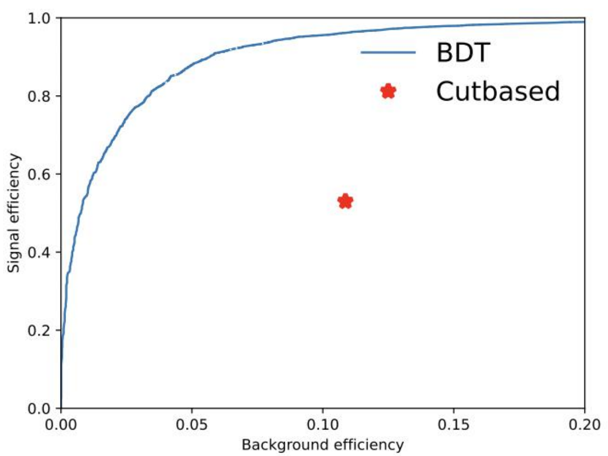
Summary of input variables for the BDT training in the two ggF categories.
$|\Delta\eta_{\mu b}^{max}|$ is the maximal $|\Delta\eta|$ between muons and bjets, while $|\Delta\eta_{\mu b}^{other}|$ is for the left muon and bjet.
$|\Delta R_{\mu b}^{min}|$ and $|\Delta R_{\mu b}^{other}|$ are defined accordingly.

Bkg rejection (BDT analysis)

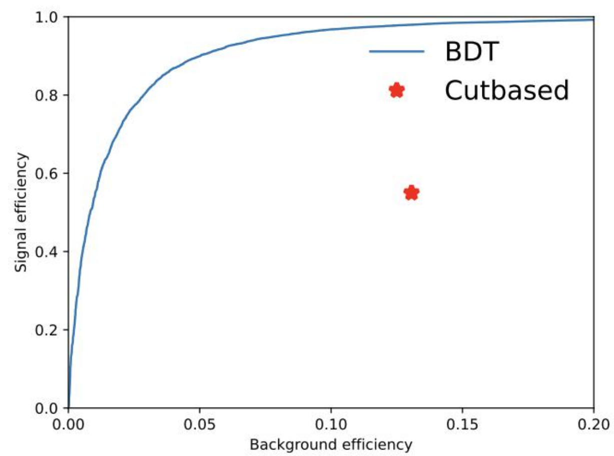




The BDT score distributions of the training and application samples for both the signal and the background



ggF SM category

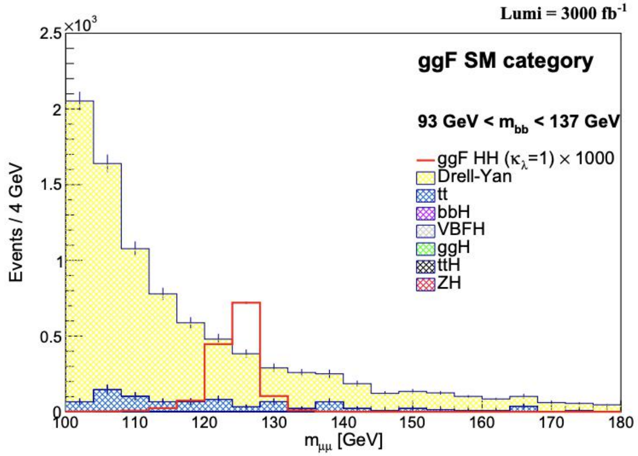


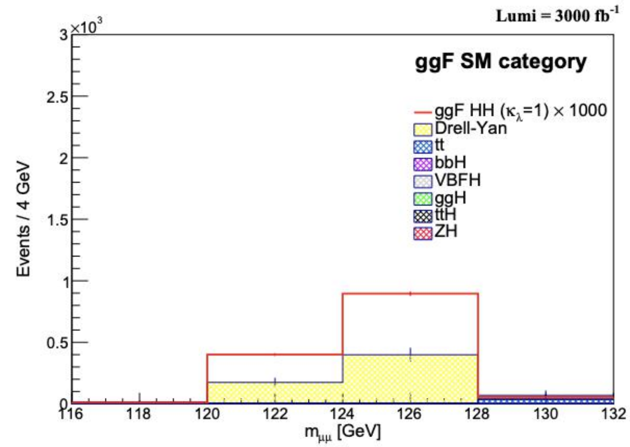
ggF BSM category

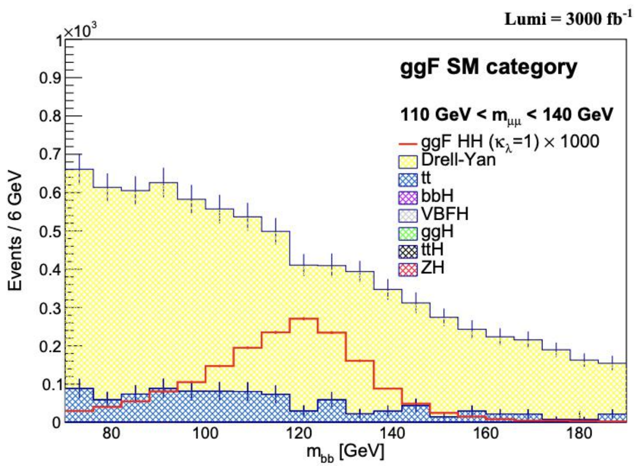
The improvement by BDT is visualized in the ROC curves.

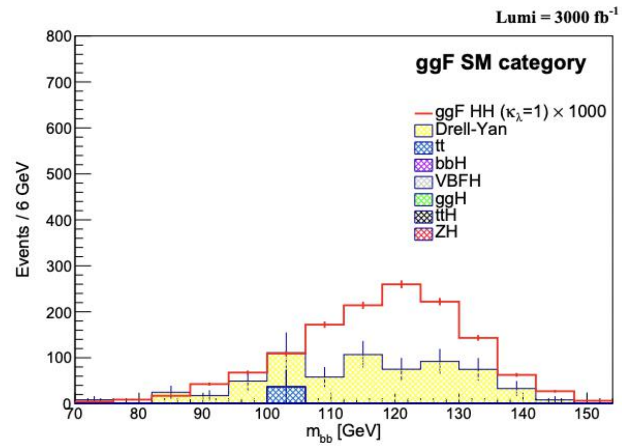
The cut-based performance is shown as the red stars for comparisons.

Fitting template



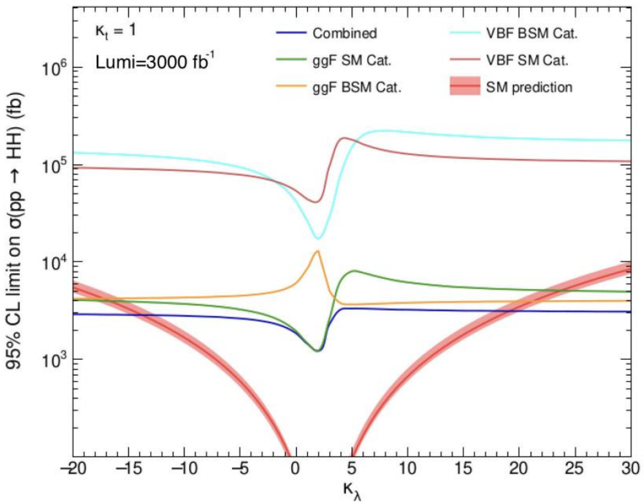


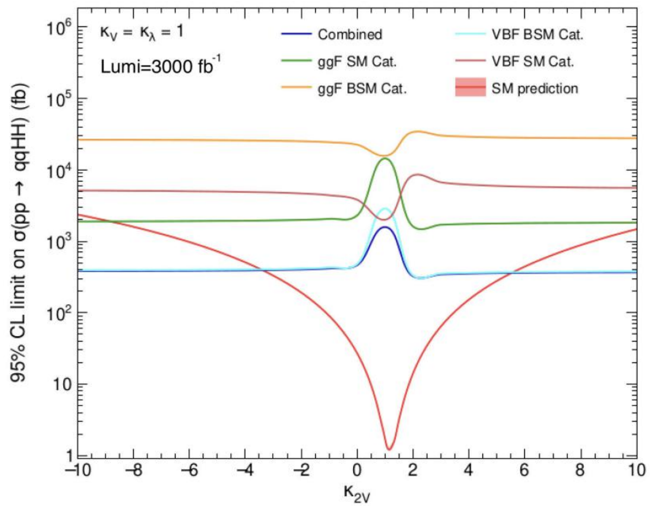




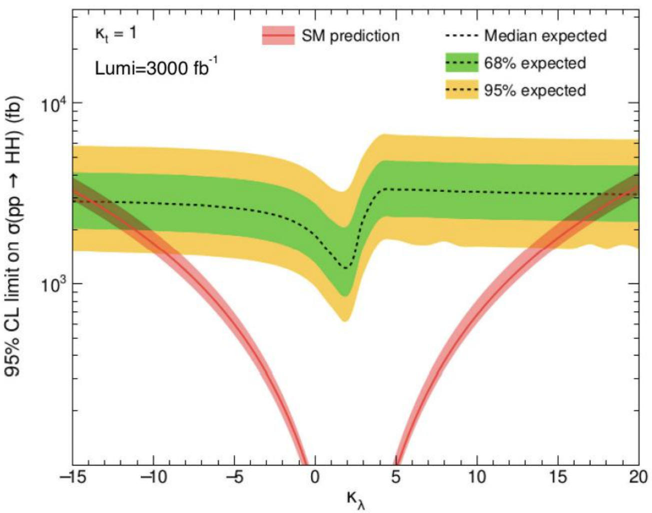
The fitting templates are the combined di-muon mass and di-bjet mass distributions, as shown in these four plots. Left for cut-based and right for BDT.
These four plots only represents ggF SM category. Other categories are not shown here.

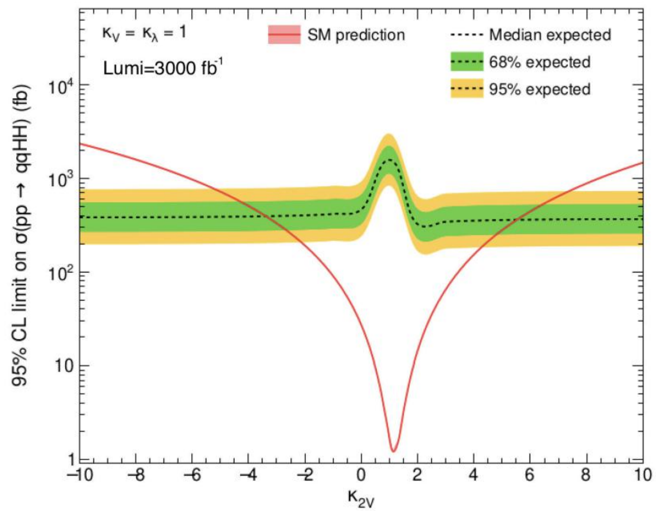
Result (Cutbased analysis)





The scan on k_{λ} is shown in upper left plots with the individual contributions from each category and the combined one, and in lower left with the combined results using the cut-based approach. The red solid curve represents the theoretical prediction. Right are corresponding the k_{2V} result.



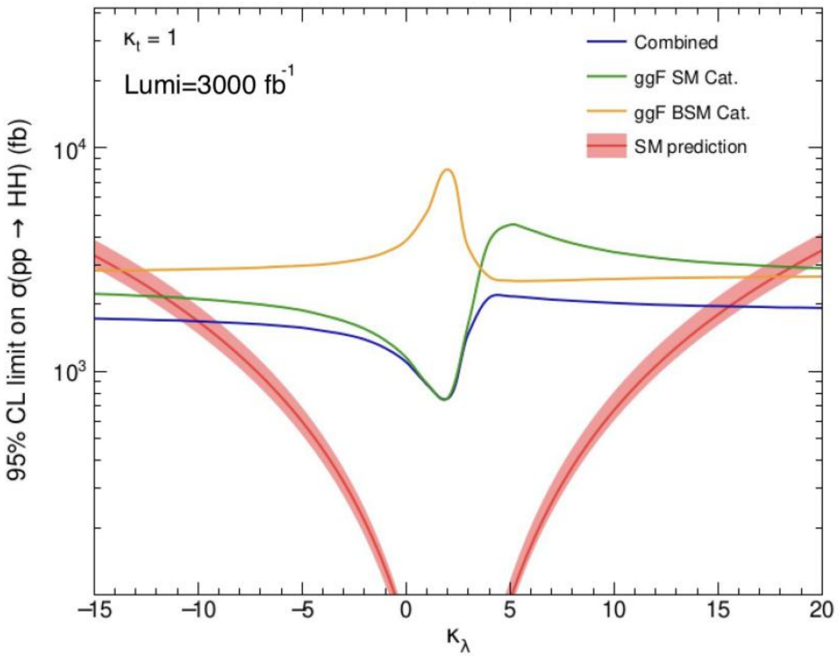


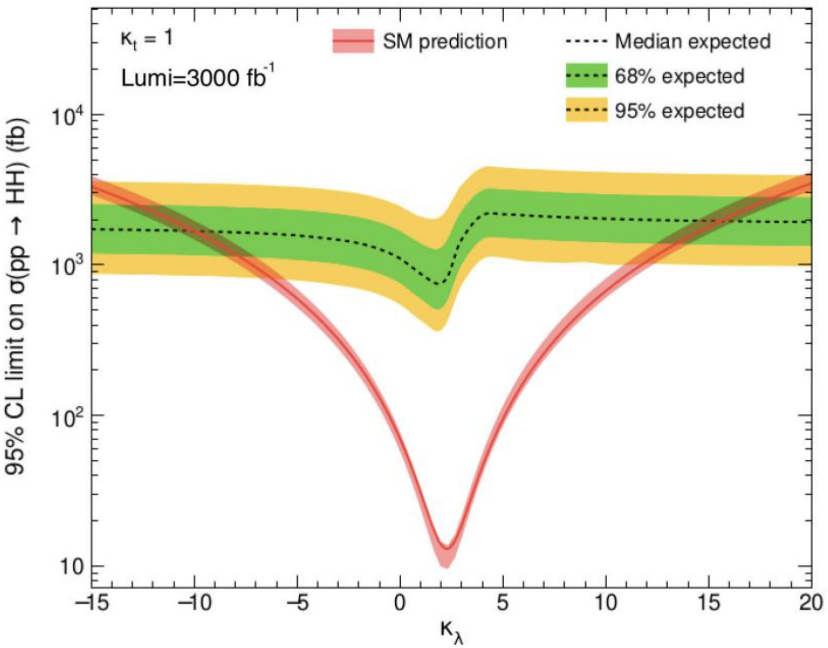
The expected exclusion is $\kappa_{\lambda} < -13.8$ and $\kappa_{\lambda} > 19.1$

The expected exclusion is $\kappa_{2V} < -3.4$ and $\kappa_{2V} > 5.5$

The expected upper limit is corresponds to 47 times the standard model prediction(Lumi=3000 fb^{-1}).

Result (BDT analysis)





The same scan on k_λ is performed using the BDT approach as shown in left for the breakdown and right for the combined only.

The expected exclusion is $\kappa_\lambda < -10.0$ and $\kappa_\lambda > 15.5$.

The expected upper limit is corresponds to 28 times the standard model prediction(Lumi=3000 fb^{-1}).

Conclusion

- We presents a comprehensive study of the Higgs boson pair production in the rare decay of $HH \rightarrow b\bar{b}\mu^+\mu^-$ with both the ggF and VBF production modes included for the first time.
- With a luminosity up to 3000 fb^{-1}, the channel $HH \rightarrow b\bar{b}\mu^+\mu^-$ can not lead to the observation of HH with the cut-based or the BDT approach.
- It is still able to contribute in a sizeable way to the HH search combination and can be sensitive to BSM enhancement given its small rate and excellent di-muon peak.

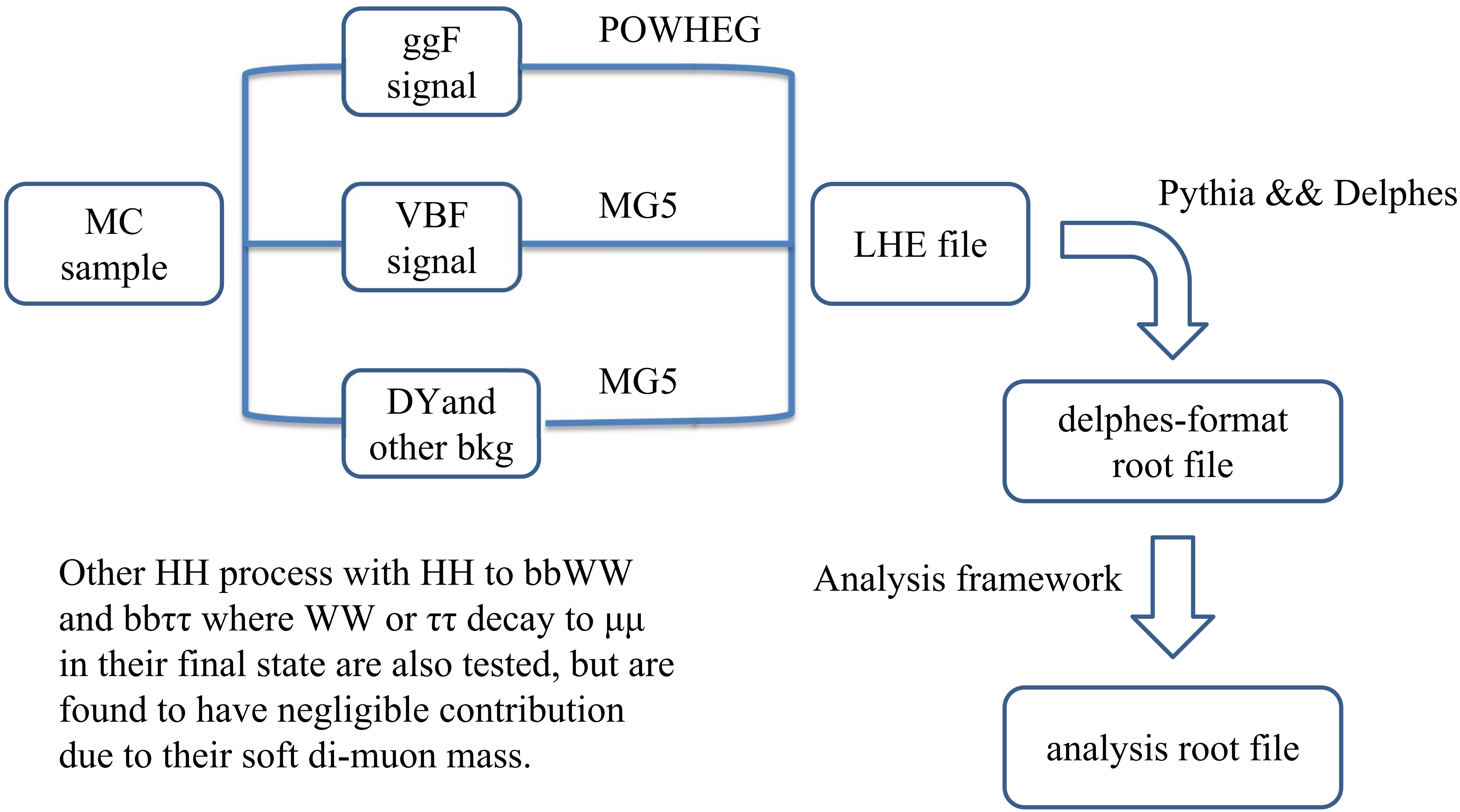
Analysis Type	300 fb^{-1}	450 fb^{-1}	3000 fb^{-1}
	ggF HH (σ/σ_{SM})		
Cut-based	152^{+87}_{-46}	123^{+70}_{-37}	$47^{+26.1}_{-14.1}$
BDT	$96^{+56}_{-29.8}$	$77^{+45}_{-23.9}$	$28^{+16.3}_{-8.8}$
	ggF+VBF HH ($\sigma/\sigma_{SM(ggF+VBF)}$)		
Cut-based	152^{+86}_{-46}	122^{+70}_{-37}	$46^{+26.1}_{-13.9}$
BDT	$96^{+56}_{-29.7}$	$77^{+45}_{-23.9}$	$28^{+16.2}_{-8.8}$
	VBF HH (σ/σ_{SM})		
Cut-based	3195^{+1440}_{-960}	2555^{+1130}_{-760}	928^{+380}_{-265}

Analysis Type	300 fb^{-1}	450 fb^{-1}	3000 fb^{-1}
	κ_λ scan		
Cut-based	(-26.9,32.2)	(-24.0,29.3)	(-13.8,19.1)
BDT	(-20.7,26.2)	(-18.3,23.8)	(-10.0,15.5)
	κ_{2V} scan		
Cut-based	(-7.6,9.8)	(-6.6,8.8)	(-3.4,5.5)

Thank you !

Back Up

MC sample

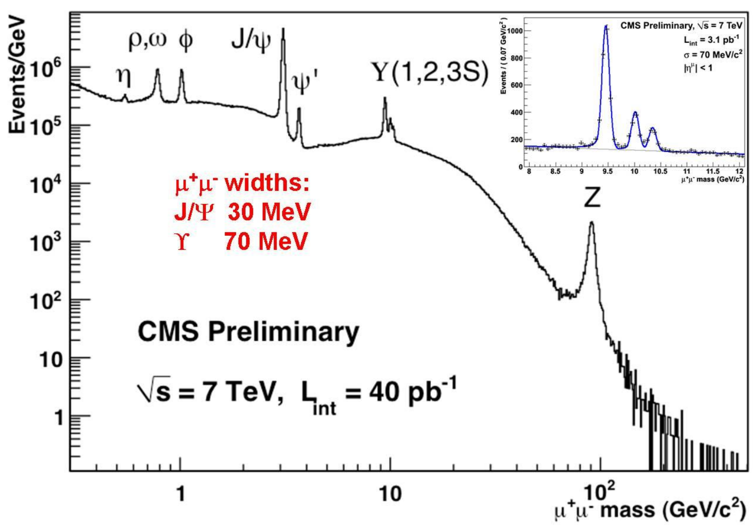


Other HH process with HH to bbWW and bbττ where WW or ττ decay to μμ in their final state are also tested, but are found to have negligible contribution due to their soft di-muon mass.

In terms of the expected sensitivity, the upper limits from the leading decay channels reach around 4 times the SM prediction on the ggF HH crosssection, while the combined results start to get close to 2 times.

The reason that we choose this decay channel is

$H \rightarrow bb$: The largest BR
$H \rightarrow \mu\mu$: The excellent resolution of $m_{\mu\mu}$



Process	$m_{l^+l^-}$ [GeV]	σ[fb]	$N^{\text{gen}}_{\text{events}}(\times 10^6)$
Drell-Yan	$[100, 150]$	5481	9.98
Drell-Yan	$[150, 200]$	384	10.0
Drell-Yan	$[200, +\infty]$	201	1.0
$t\bar{t}$	—	4864	2.0
ggH	—	82836	1.0
VBFH	—	4058	1.0
ZH	—	1775	1.0
ttH	—	836	1.0
bbH	—	638	1.0
ggF signal			
$\kappa_\lambda = -5$	—	599	0.55
$\kappa_\lambda = 0$	—	70	0.55
$\kappa_\lambda = 1$	—	31	0.55
$\kappa_\lambda = 2.4$	—	13	0.55
$\kappa_\lambda = 5$	—	95	0.55
$\kappa_\lambda = 10$	—	672	0.55
$\kappa_\lambda = 20$	—	3486	0.55
VBF signal			
$\kappa_{2V} = -10$	—	2365	0.50
$\kappa_{2V} = -5$	—	722	0.50
$\kappa_{2V} = 0$	—	27	0.50
$\kappa_{2V} = 1$	—	1.73	0.50
$\kappa_{2V} = 2$	—	14.2	0.50
$\kappa_{2V} = 5$	—	279	0.50
$\kappa_{2V} = 10$	—	1479	0.50

TABLE I. Summary of Monte Carlo samples.

HL-LHC is expected to have 23 events that HH to bbμμ.

Some parameters for BDT

The training setup includes 2500 trees, the tree depth of 3 and a learning rate of 0.08 (0.1) for ggF SM category (ggF BSM category).
The MC samples are splitted into 64%, 16% and 20% for training, testing and application.

The signal samples used in the ggF SM category include $\kappa\lambda = 1$.
The signal samples used in the ggF BSM category include $\kappa\lambda = 5, 10, 20$.
Both the DY and tt processes are used as the background in the training.

We use only MC sample to do the search, not involving the response of the detector.
So it doesn't contain the system uncertainty.