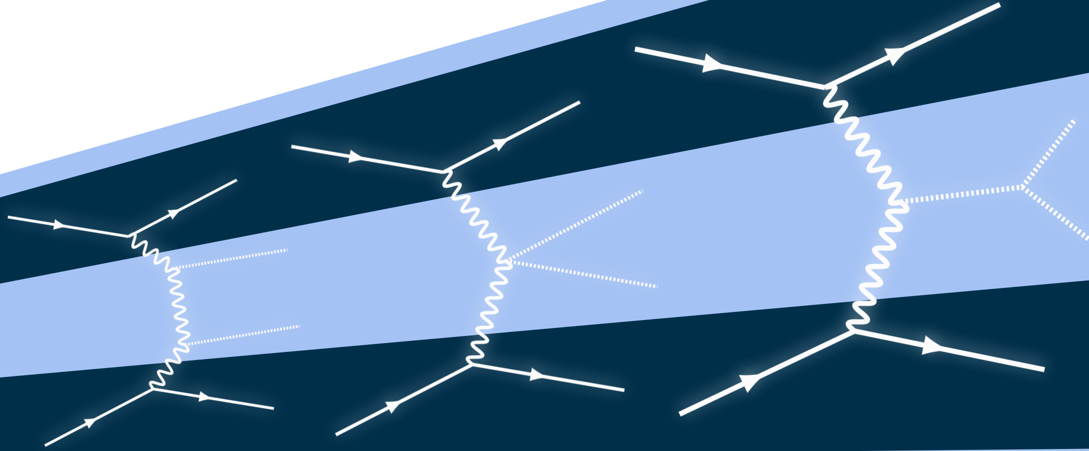
An Initial Look at a VBF HH κ_λ Scan

James Grundy
(University of Oxford)

HH VBF Taskforce Meeting
05/10/20



Introduction

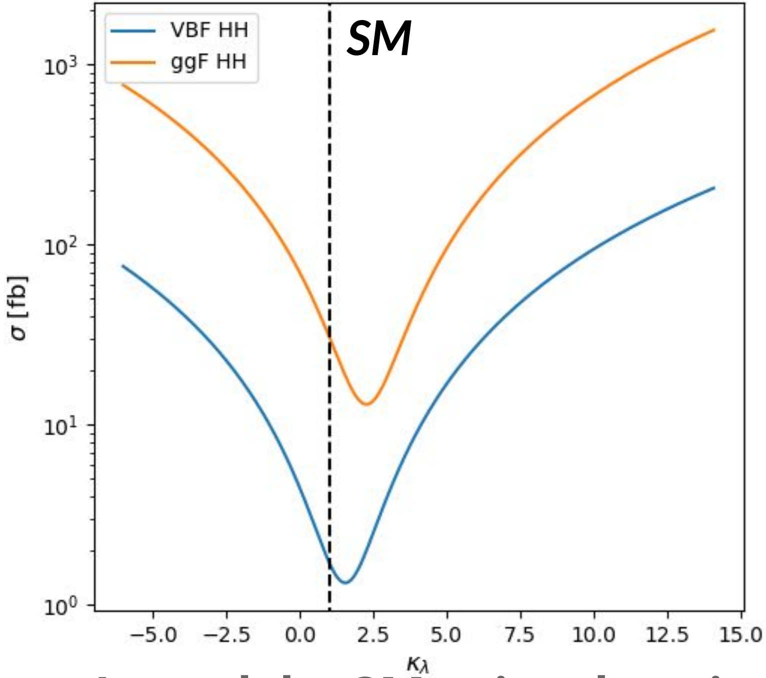
There is ongoing discussion regarding the result(s) the VBF HH analyses should aim for.

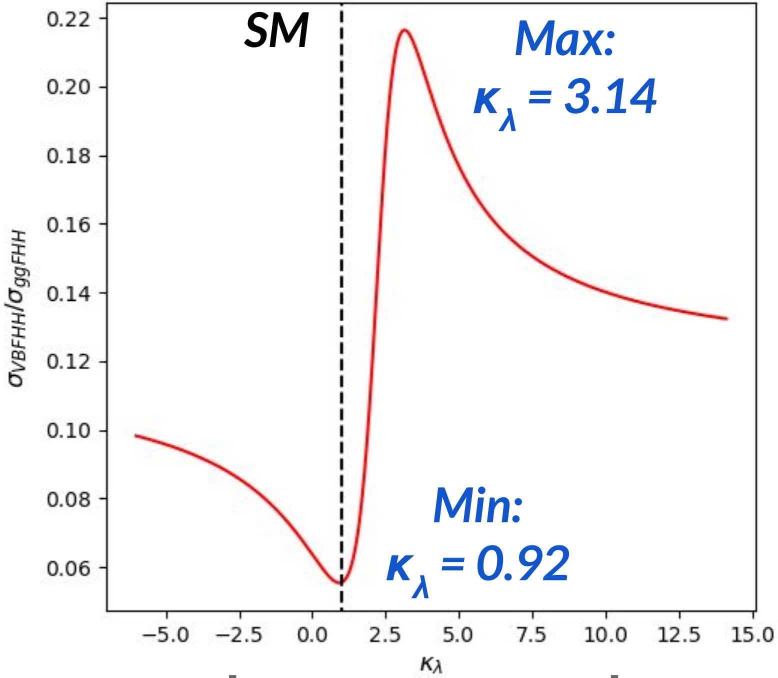
Suggested aims for VBF HH→4b analysis team:

- **Include VBF in main (ggF) non-resonant κ_λ scan.**
- **Exclude $c_{2V} = 0$.**
- **Generate first 2D κ_λ-κ_{2V} scan (timescale of ~1-2 years).**

Q: are we confident VBF will improve main non-resonant κ_λ scan? Today, I'll go through some studies that will help inform a judgement on this.

Comparing ggF and VBF HH σ





Around the SM point, there is the biggest gap between σ_{VBF} and σ_{ggF}.
VBF *looks* more competitive at κ_λ points > 1.

Accounting for VHH Contamination

VBF HH MC samples also contain VHH events → LO σ on AMI is VBF+VHH → require correction to scale to just VBF σ.

Correction factor is highlighted below. Measured the VBF fraction in VBF HH EXOT8 MC16a (thanks to Tulin and Chris!).

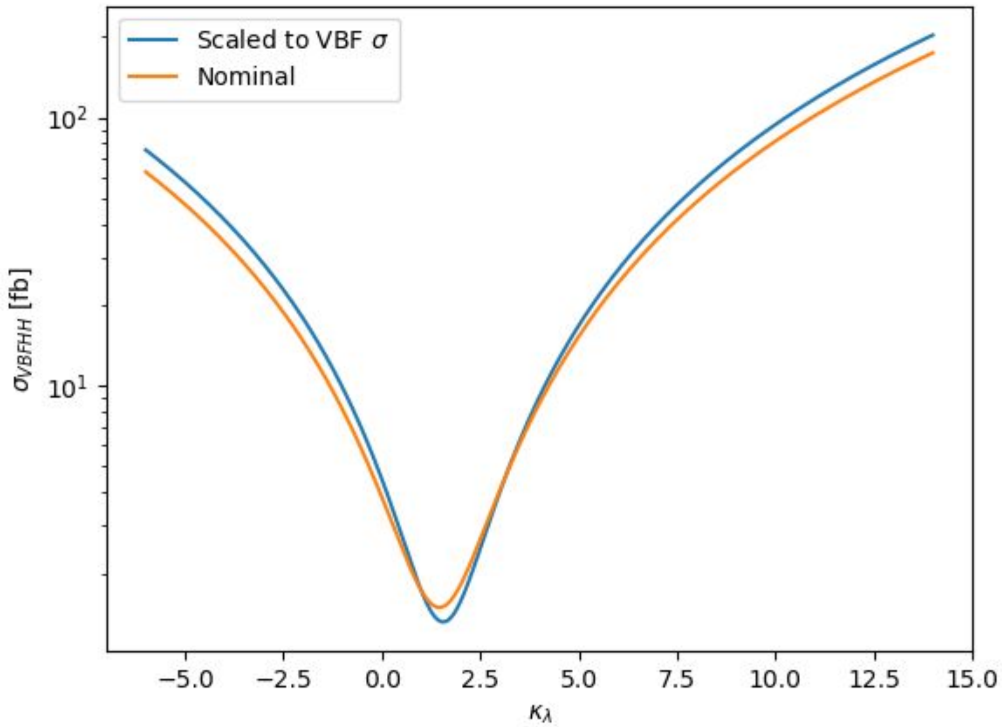
$$\sigma_{N3LO}^{\lambda=x} = \sigma_{LO}^{\lambda=x} \frac{\sigma_{N3LO}^{\lambda=1}}{\sigma_{LO}^{\lambda=1}} \boxed{\frac{(VBF\ fraction\ at\ \lambda{=}x)}{(VBF\ fraction\ at\ \lambda{=}1)}}$$

Accounting for VHH Contamination

VBF HHbbb EXOT8 MC16a

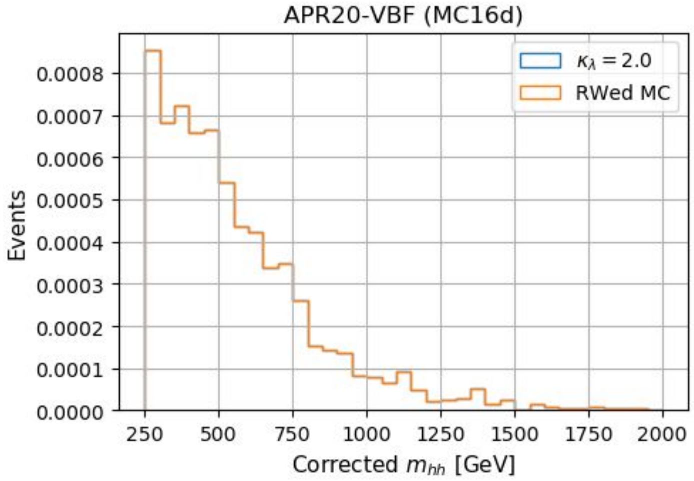
κ_λ	VBF Fraction	VHH Fraction
0	0.95494	0.04506
1	0.81864	0.18136
2	0.70975	0.29025
10	0.94405	0.05595

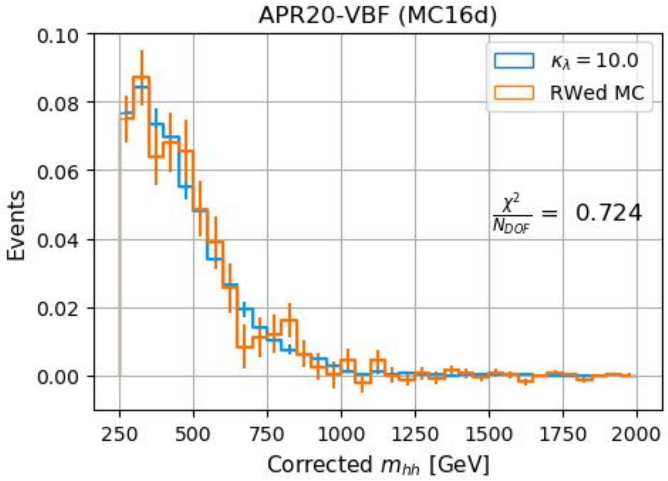
Good agreement w/ bbγγ & bbττ



To study all κ_λ points for VBF HH, implemented MC combination and re-weightinging procedure used for c_{2V} in previous VBF HH→4b analysis.

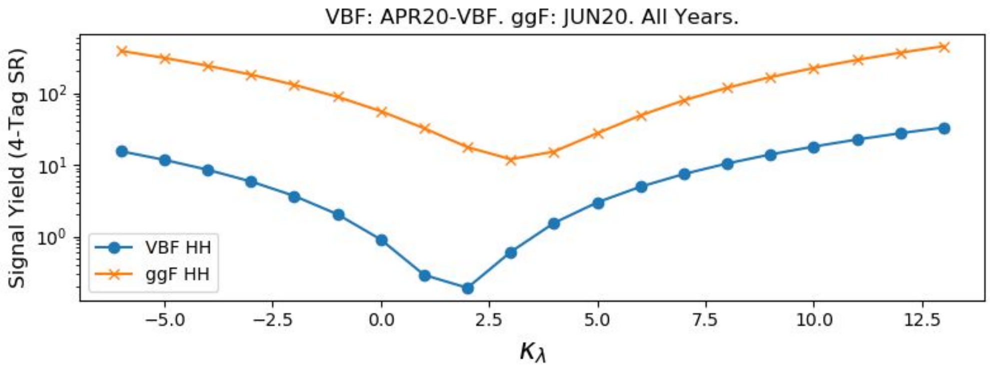
$$|A(c_V, c_{2V}, \lambda)|^2 = \left(1 - \tfrac{3}{2}\lambda + \tfrac{1}{2}\lambda^2\right)|A(1,1,0)|^2 + \left(2\lambda - \lambda^2\right)|A(1,1,1)|^2 + \tfrac{1}{2}\left(\lambda^2 - \lambda\right)|A(1,1,2)|^2$$

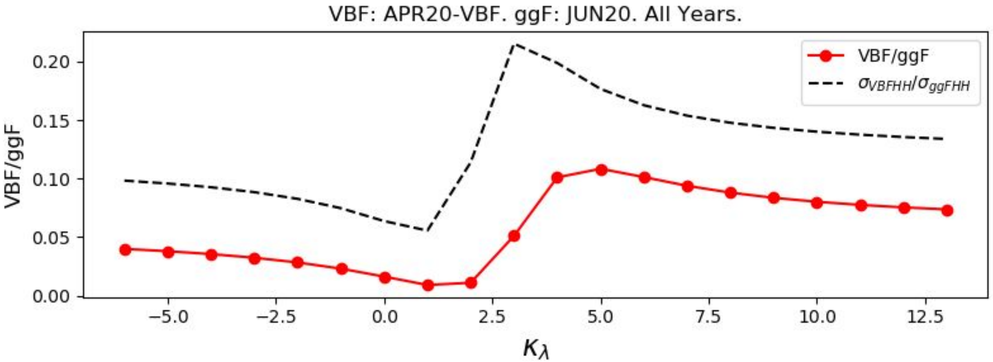




For ggF HH, I used a set of weights derived in the previous ggF HH→4b analysis (thanks to Nicole!).

Comparing Signal Yields

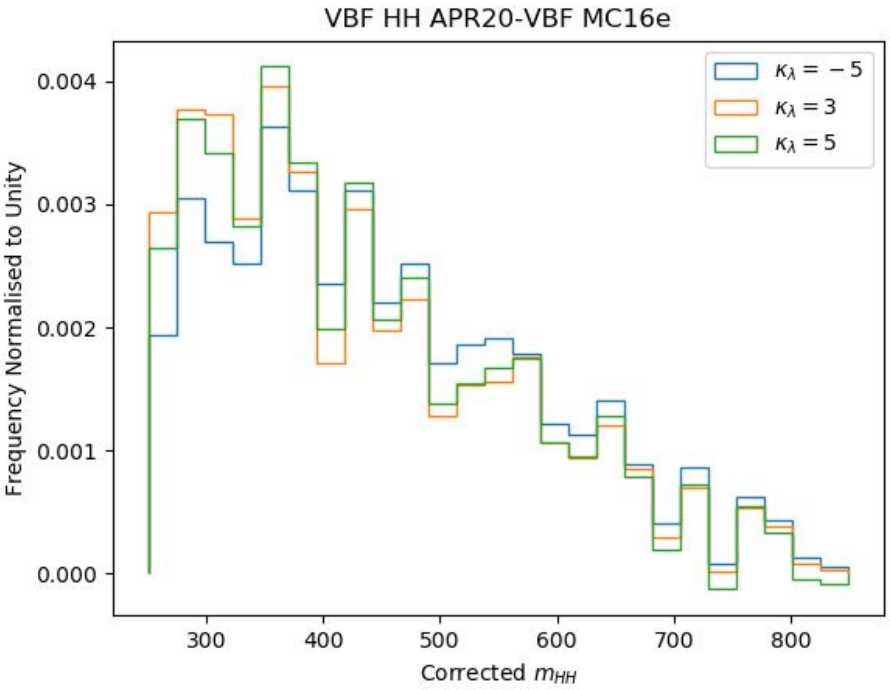


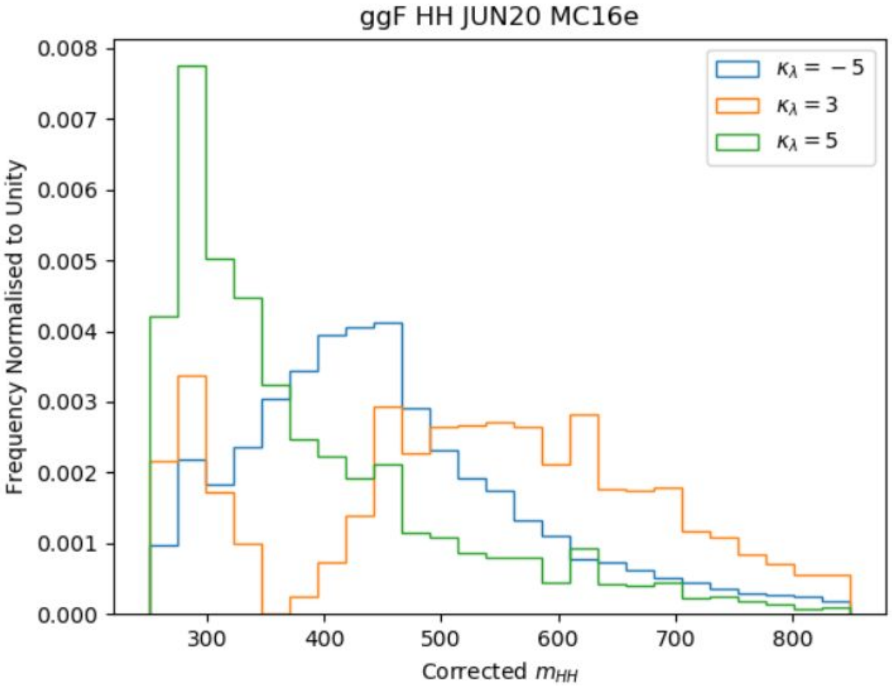


Main Analysis NNTs: JUN20.
VBF Analysis NNTs: APR20-VBF (has issue w/ MC-MC SFs).

Q: are the BSM VBF events in a sensitive region of phase space?

Ratio of signal yields does reflect σ ratio, but peak is lost.





VBF mHH shape ~stable w/ BSM κλ, change in signal yield ratio is driven by ggF.

+ κλ → soft ggF → decreased acceptance → VBF is more competitive.

- κλ → hard ggF → increased acceptance → VBF is less competitive.

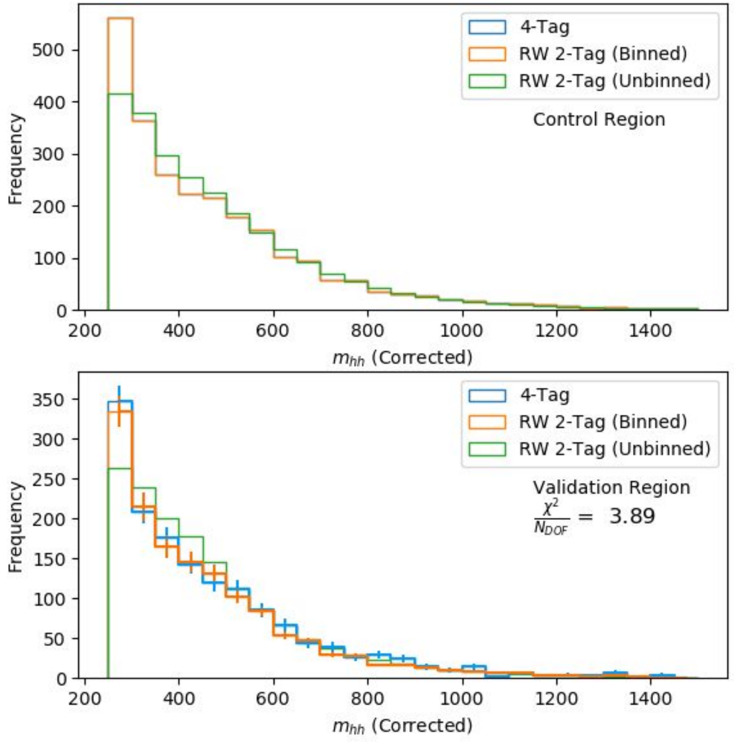
κλ = 3.0 → majority hard ggF → increased acceptance → VBF is less competitive.

Approximating Backgrounds

Generate background estimate by RWing 2-tag SR data using binned SFs derived in CR.

□^{2} value is driven by highest mHH bins. Range 0-1000 has □^{2} ~= 1 (back up) → statistical errors in 4-tag are very large and go some way to covering shape difference.

APR20-VBF (2016-2018 Data)



Measuring Sensitivity

Using corrected mHH distributions for 4-Tag SR MC and the aforementioned background estimate, I calculated:

- **Quadrature sum of binned significances, Z, (*paper here*):**

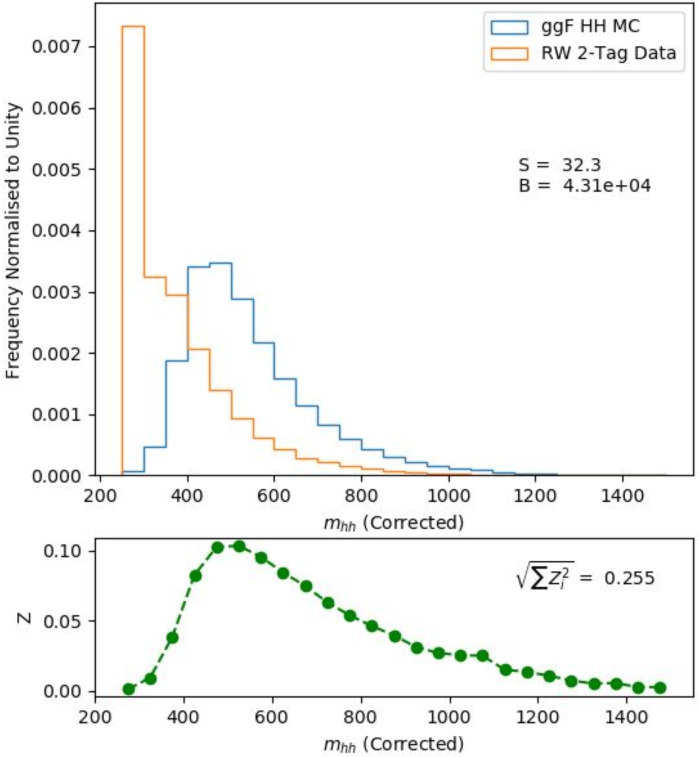
$$Z = \begin{cases} +\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1 + \frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{if } n \geqslant b \\ -\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1 + \frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{if } n < b. \end{cases}$$

- **Approximation of the limit on μ (*Bill's Thesis (6.2.4)* + b/u).**

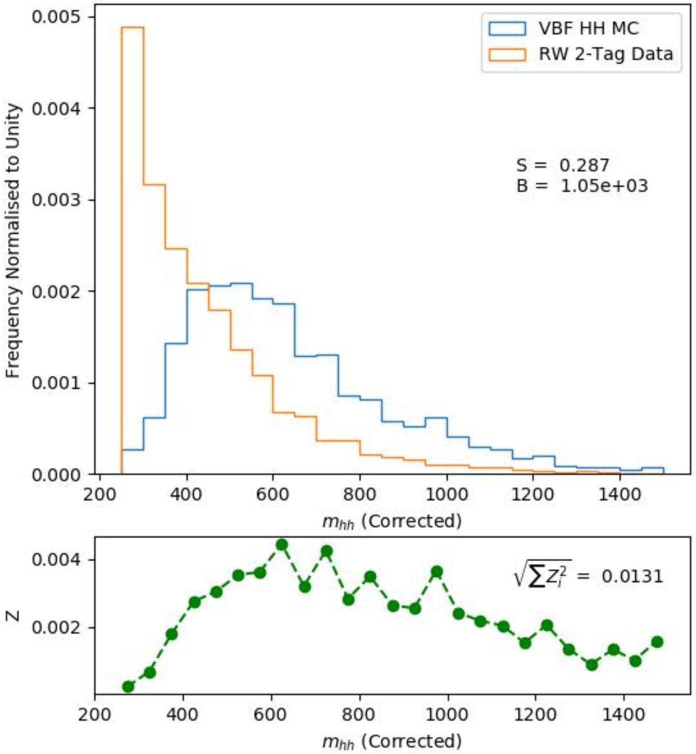
Measuring Sensitivity

Statistical Error Only

ggF SM



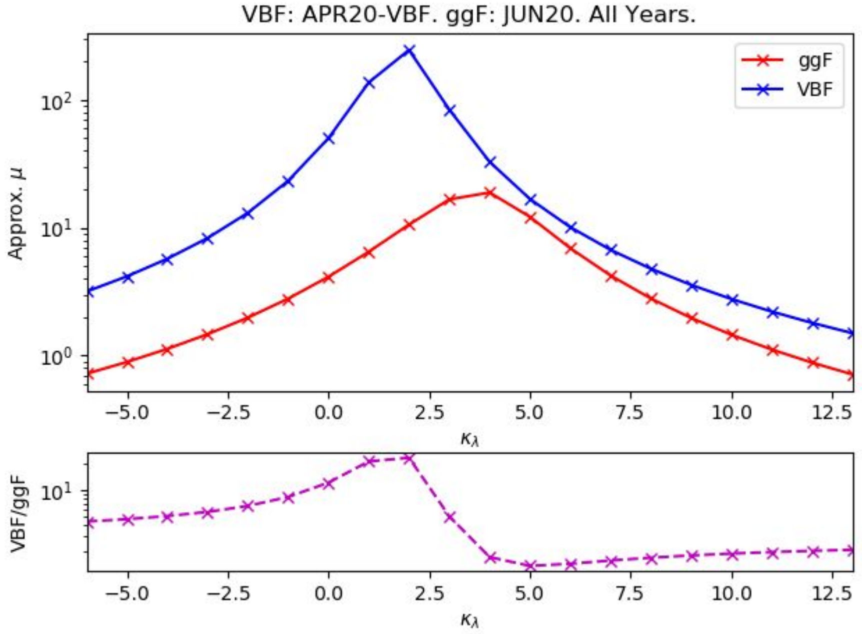
VBF SM



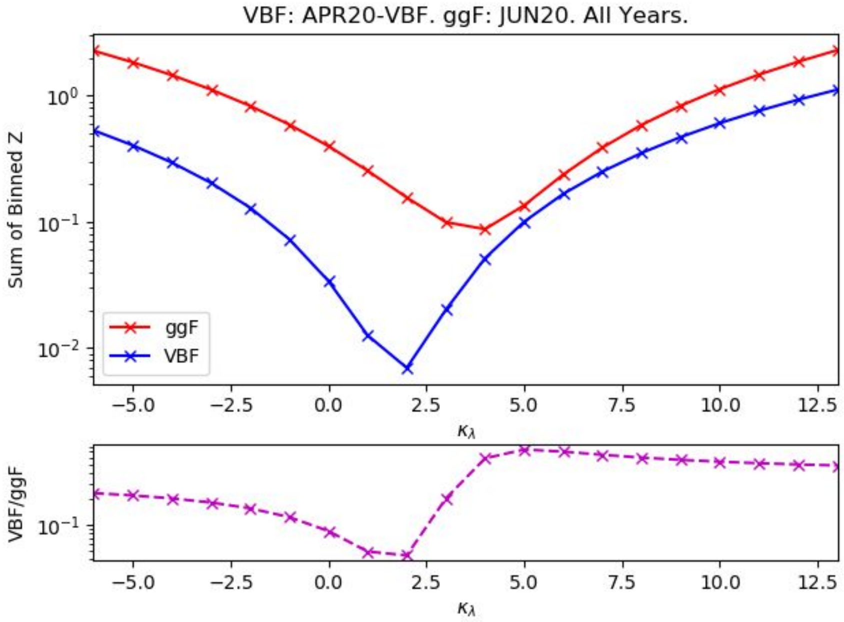
Comparing Sensitivity

Statistical Error Only

Approx. μ



Sum of Binned Z

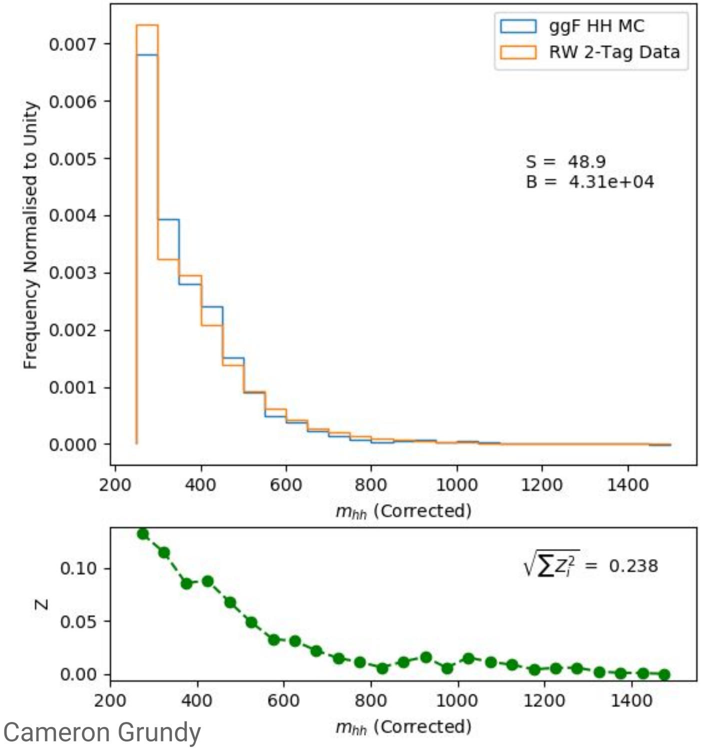


At higher +κλ, our sensitivity to VBF becomes a lot closer to that of ggF.

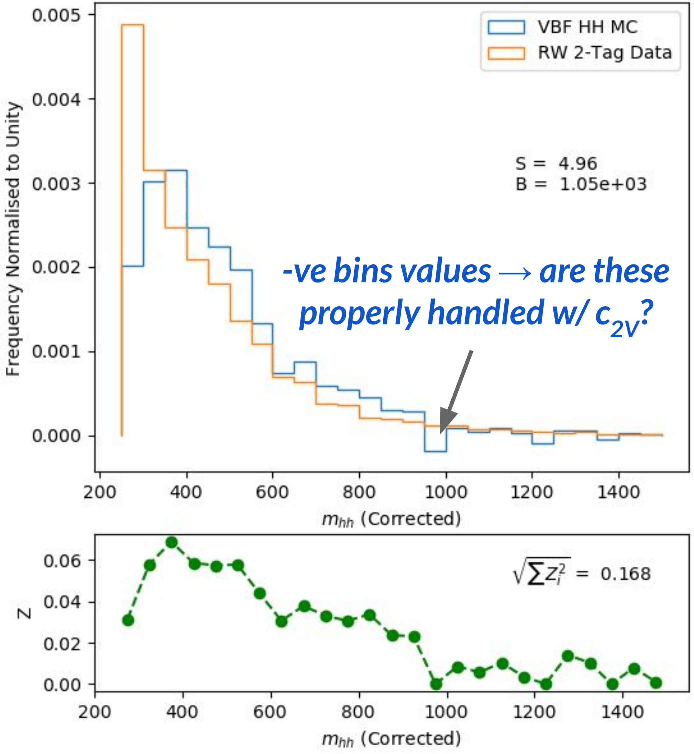
Comparing Sensitivity

Statistical Error Only

ggF $\kappa_\lambda = 6.0$



VBF $\kappa_\lambda = 6.0$



-ve bins values → are these properly handled w/ c_{2V}?

Including Uncertainties

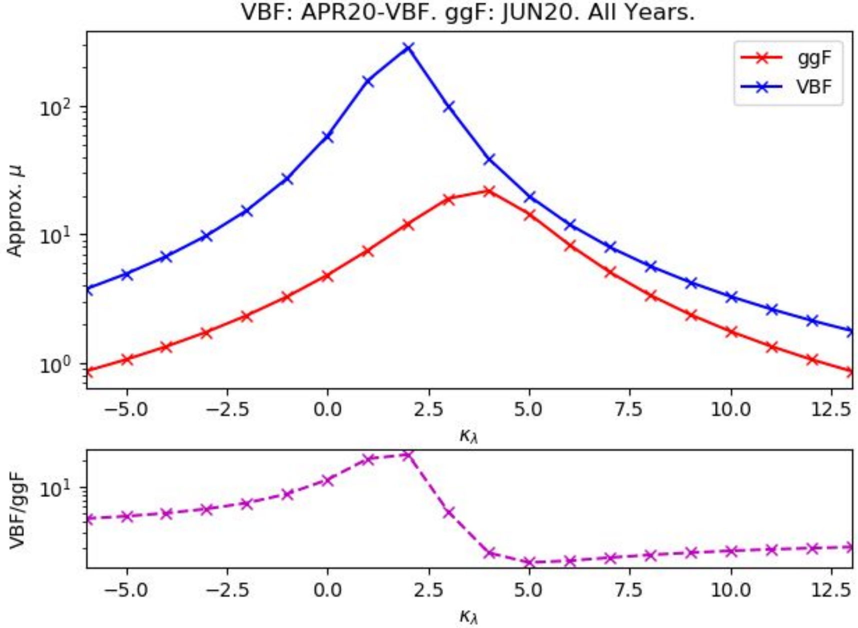
Low statistics in VBF analysis CR → uncertainty on bkg est SFs. To account for this, approximated the bkg est error as the difference between the nominal bkg est and the bkg est generated with SFs at their upper/lower error bounds.

	VBF Analysis		ggF Analysis	
	σ_B/B		σ_B/B	
Corrected mHH Bin	**Stat. Only**	**Stat+Syst**	**Stat. Only**	**Stat+Syst**
1	0.007	0.04	0.001	0.005
10	0.021	0.13	0.006	0.025
19	0.060	0.34	0.022	0.15

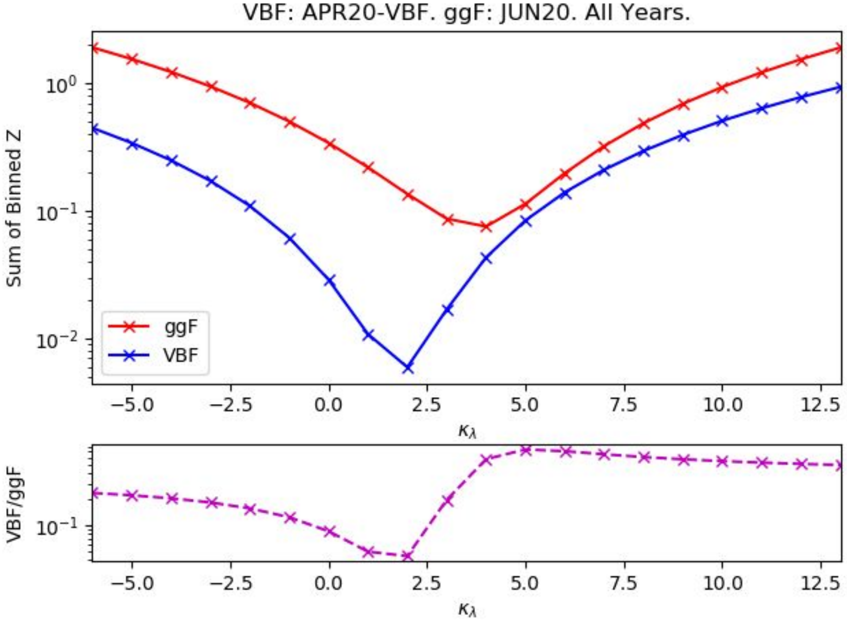
Comparing Sensitivity

Statistical Error + Bkg Est SF Uncertainty

Approx. μ



Sum of Binned Z



Including Uncertainties

ggF analysis' largest uncertainty is the shape+scale systematic on the background. Accounted with a proxy error, which was simply a constant % of the bkg est.

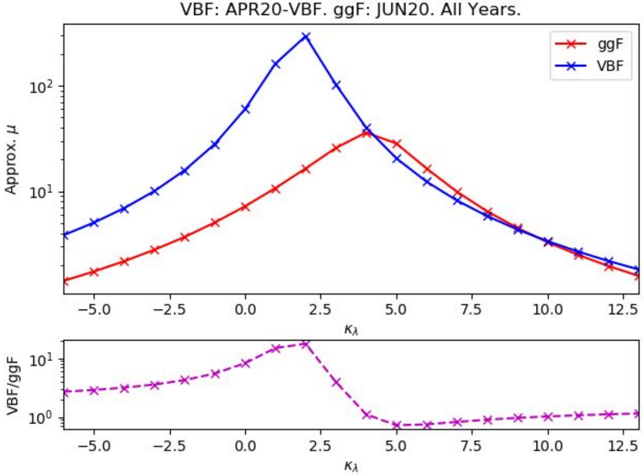
	VBF Analysis	**ggF Analysis**
Corrected mHH Bin	σ_B/B **(Stat+Syst)**	σ_B/B **(Stat+Syst)**
1	0.04	0.005
10	0.13	0.025
19	0.34	0.15

Existing error is much larger in VBF analysis than ggF → the percentage increase on overall VBF uncertainty should be less than for ggF.

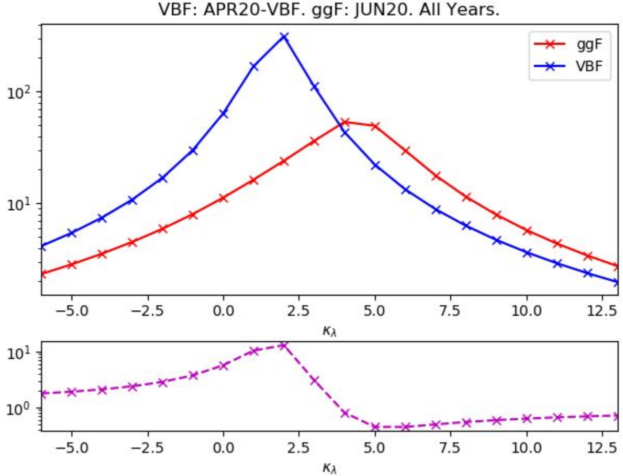
Comparing Sensitivity

Explored using a range of background systematics correlated across bins.

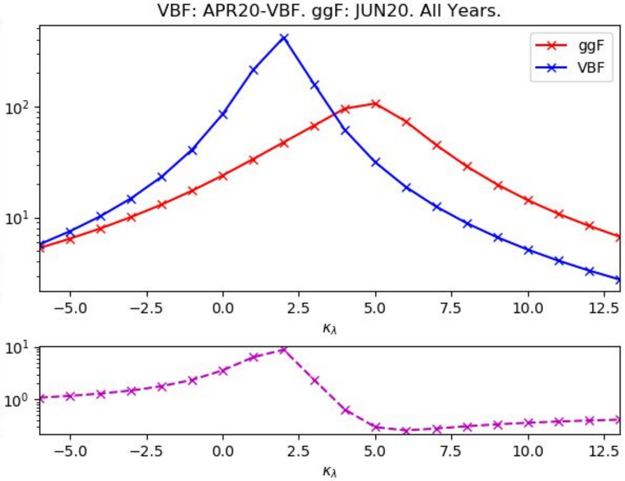
1% bkg syst



2% bkg syst

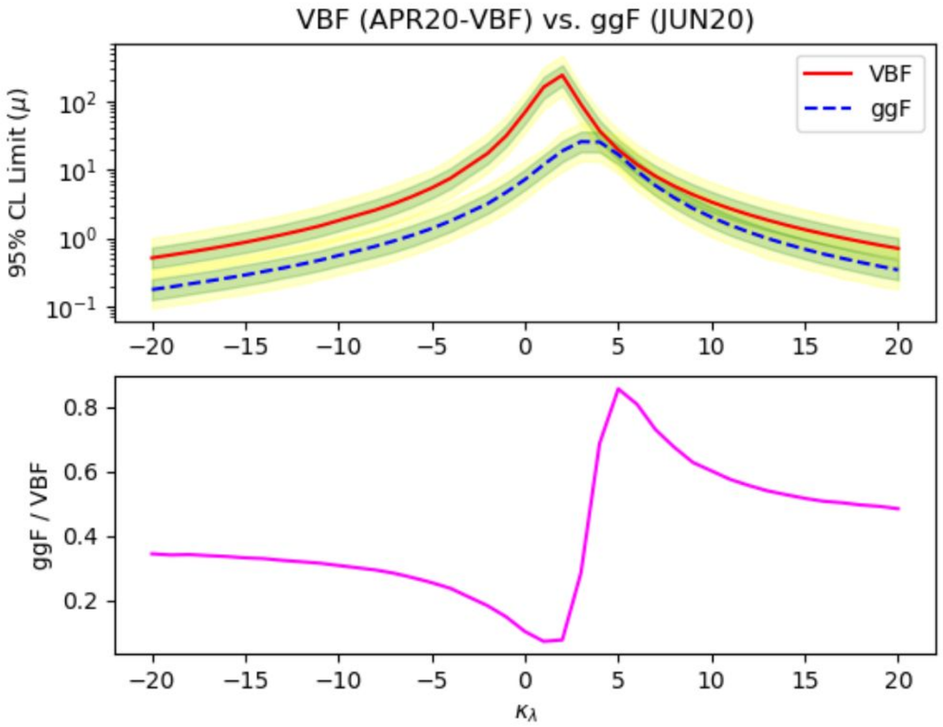


5% bkg syst



The above result indicates that our sensitivity to VBF *may* be equal (or better) than ggF when accounting for background systematics for certain κ_λ values.

Comparing 95% CL on μ



κ_λ	Ratio Value
1.0	0.07
5.0	0.86
3.0-18.0	> 0.5

ATLAS' best current exclusion limit is -2.3 < κ_λ < 10.3 (link) → VBF improves sensitivity at unexcluded kl values.

Including background systematics (1 NP for bkg shape systematic not 2 HT split NPs).

Moving Forward

What's the next step? Study a ggF-VBF Combination!

Sadly, that's not possible with pyhf...

pyhf doesn't support models w/ two POIs → must lobby them to include this!

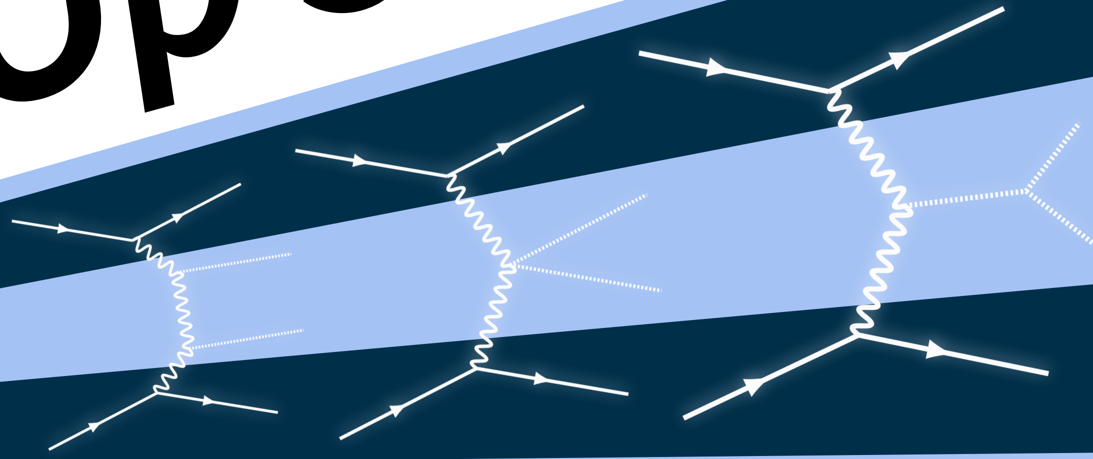
Summary and Outlook

- **Including VBF HH will not improve sensitivity at the SM point for the current HH→bbbb analysis.**
- **The inclusion of VBF HH shows significant potential to improve sensitivity for BSM points, in particular for points above κ_λ=1.**
- **Results of a ggF-VBF combination are required before a definitive statement can be made, but results so far indicate VBF HH should be included in the HH→bbbb analysis κ_λ scan.**

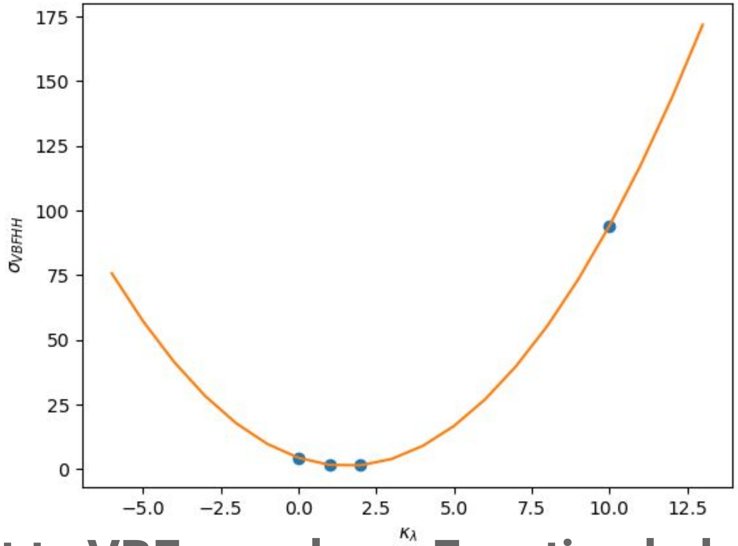
Plan going forward:

- **Explore ggF-VBF combinations.**

Back-Up Slides



$$\sigma^{\lambda=x}_{N3LO} = \sigma^{\lambda=x}_{LO} \times (VBF\ fraction\ at\ \lambda = x) \times \frac{\sigma^{\lambda=1}_{N3LO}}{\sigma^{\lambda=1}_{LO} \times (VBF\ fraction\ at\ \lambda=1)}$$



Fit to VBF xs values. Equation below.

VBF theoretical σ curve: fit quadratic polynomial to VBF HH σ (above) for four available BSM κ_λ samples.

ggF theoretical σ curve: quadratic function provided by LHCHXSWG (link).

1.302308200044769x^2 + -4.058979734141353x + 4.479354054351827

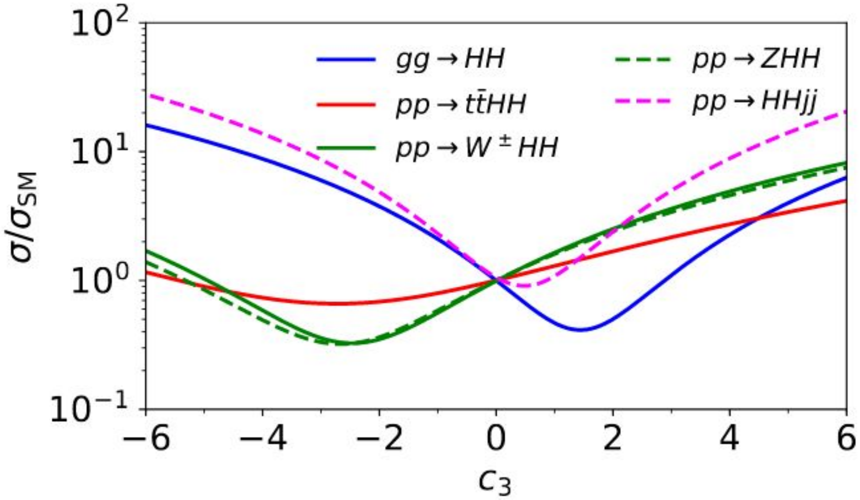
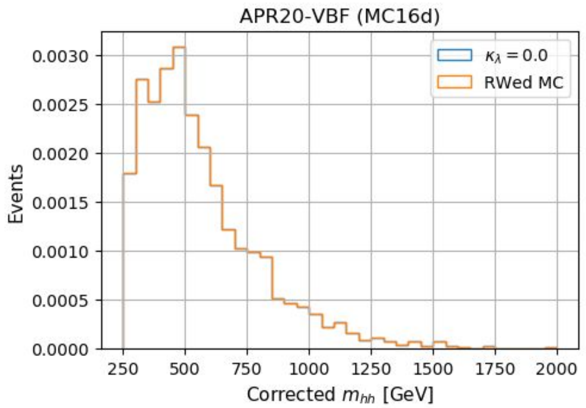
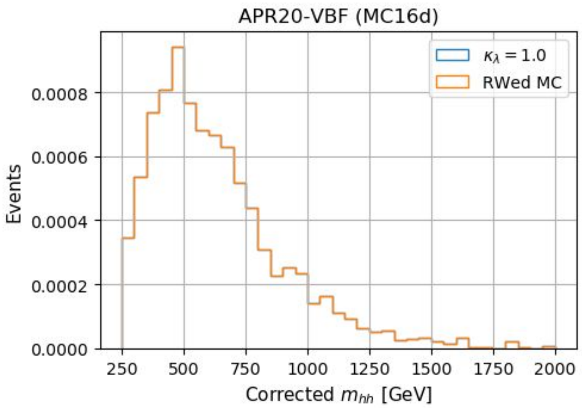
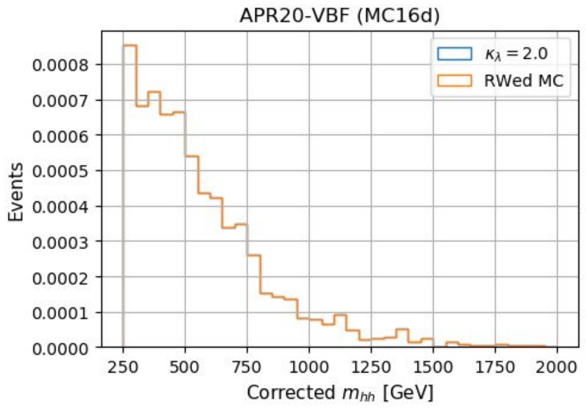
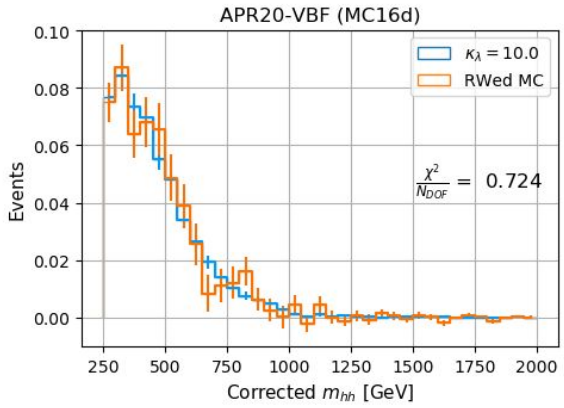


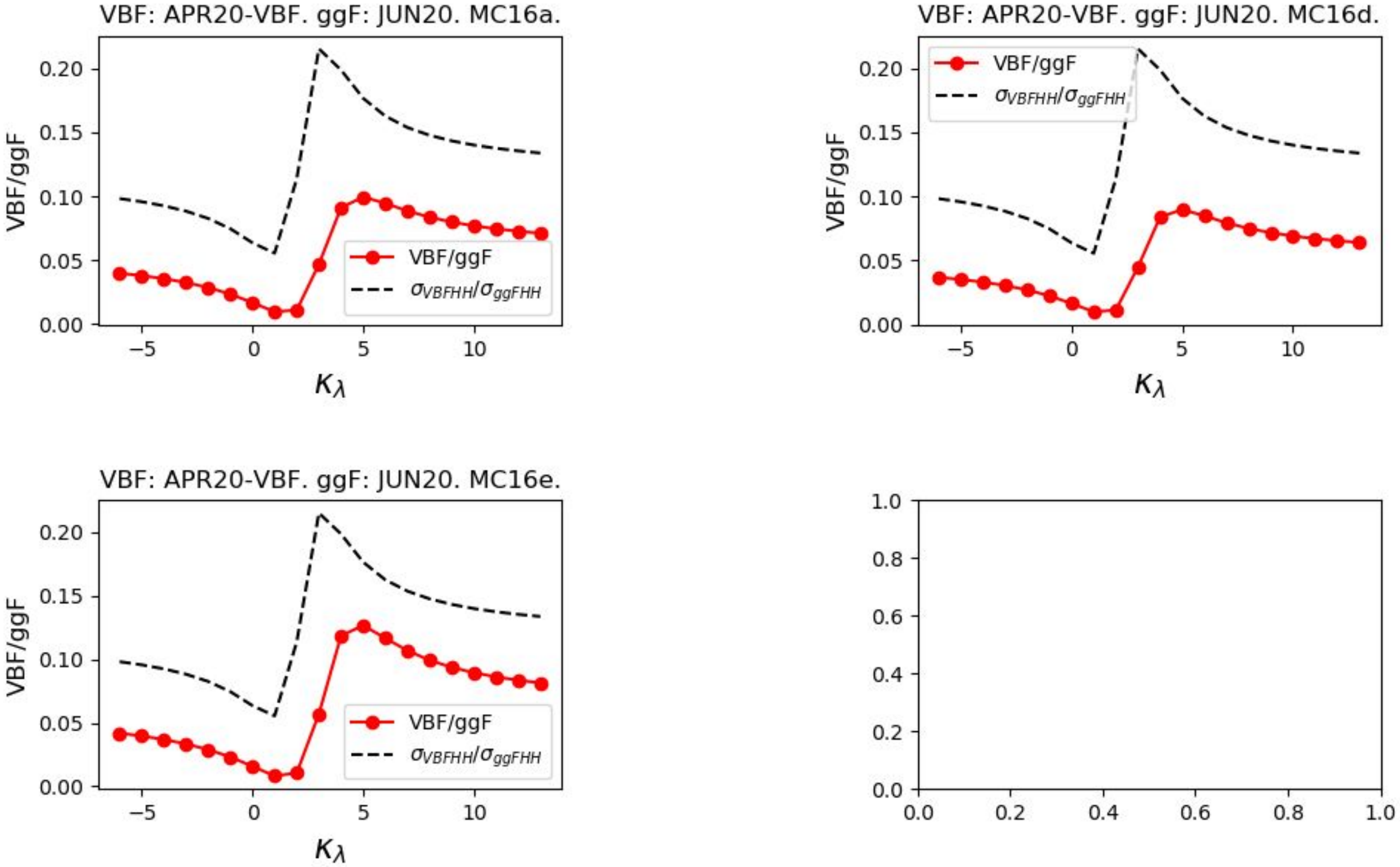
FIG. 2: The scaling of the cross section with modifications to $\lambda = \lambda_{SM}(1 + c_3)$ of the leading di-Higgs production channels. The VHH channels are at this level the most sensitive to small positive modifications to λ.


APR20-VBF (MC16d)
$\kappa_\lambda = 0.0$
RWed MC
Events
Corrected m_{hh} [GeV]

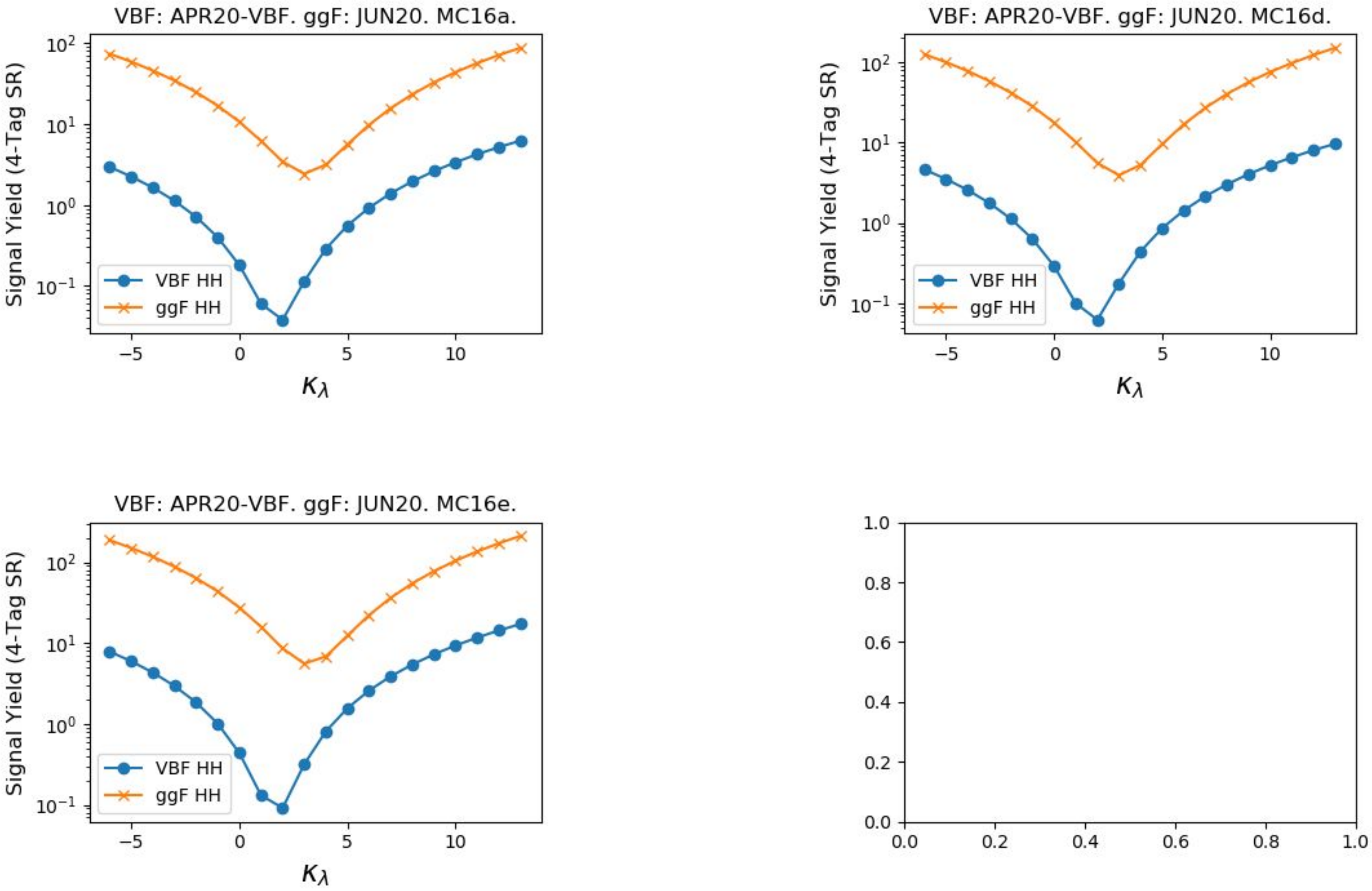

APR20-VBF (MC16d)
$\kappa_\lambda = 1.0$
RWed MC
Events
Corrected m_{hh} [GeV]

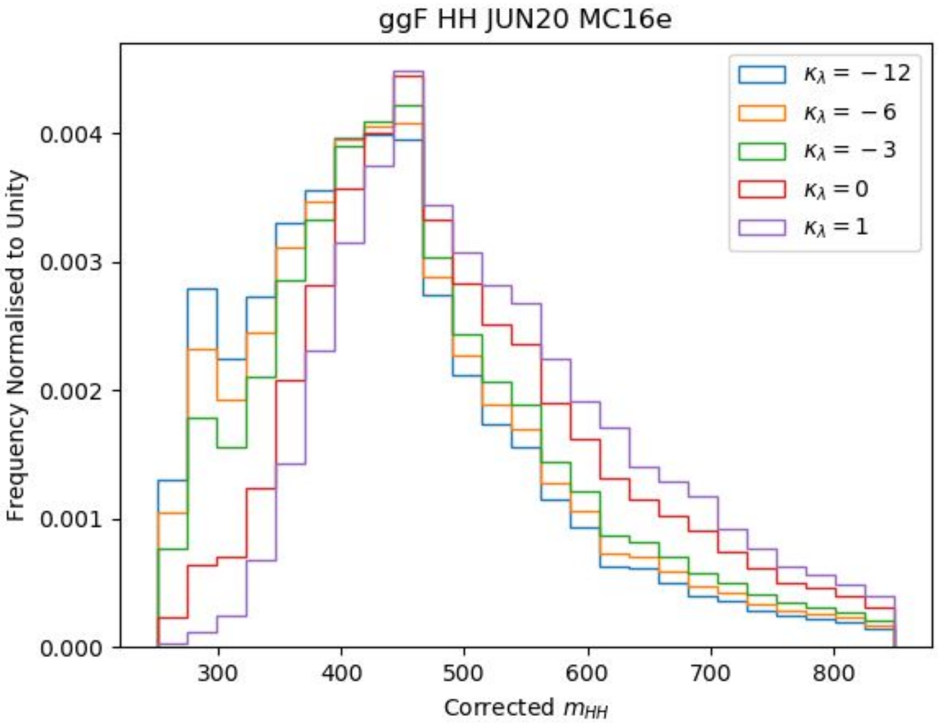

APR20-VBF (MC16d)
$\kappa_\lambda = 2.0$
RWed MC
Events
Corrected m_{hh} [GeV]

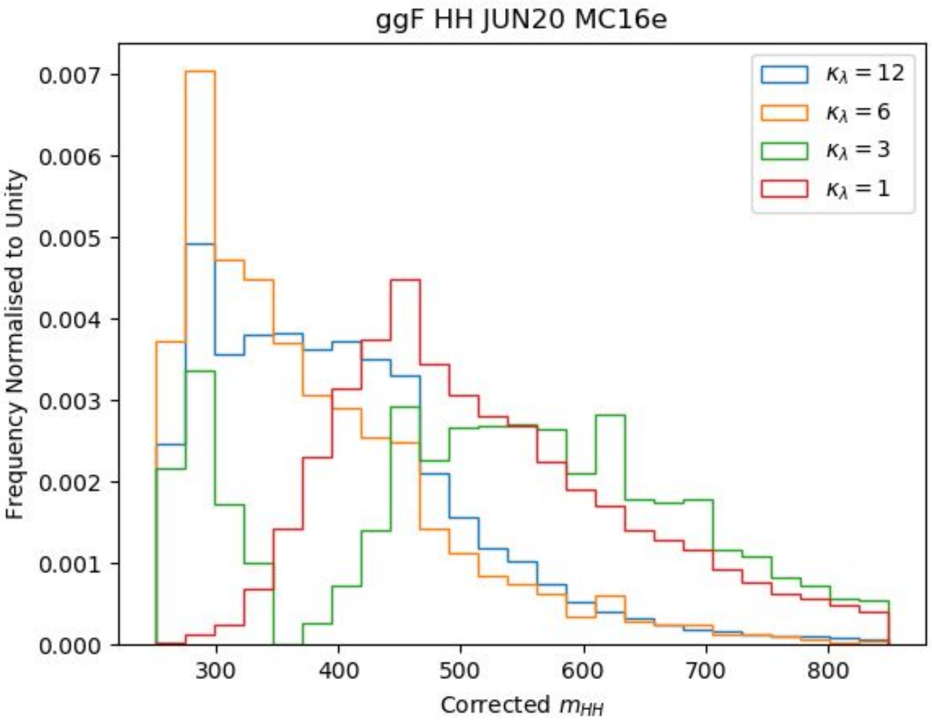

APR20-VBF (MC16d)
$\kappa_\lambda = 10.0$
RWed MC
$\frac{\chi^2}{N_{DOF}} = 0.724$
Events
Corrected m_{hh} [GeV]

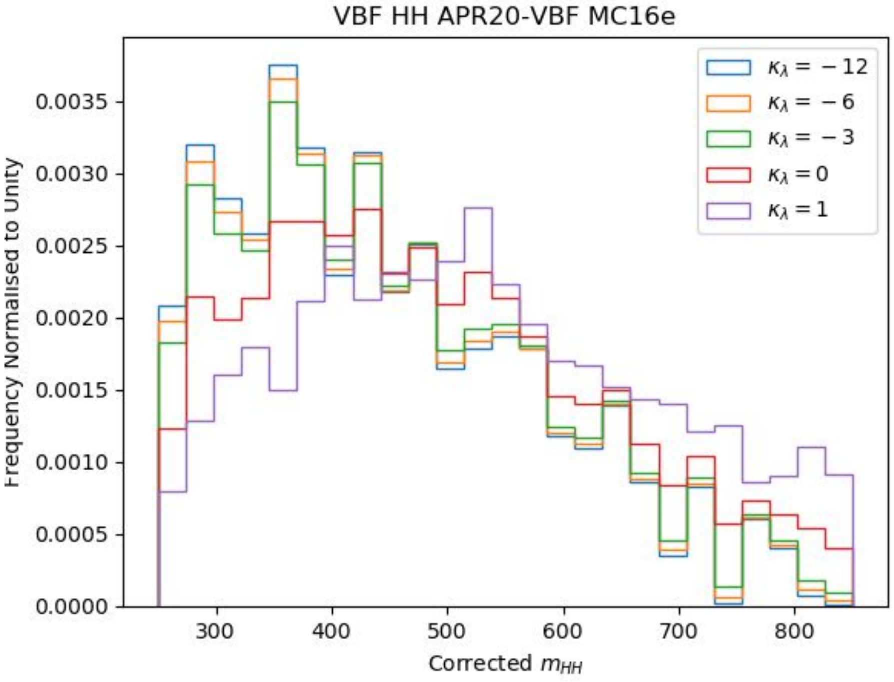


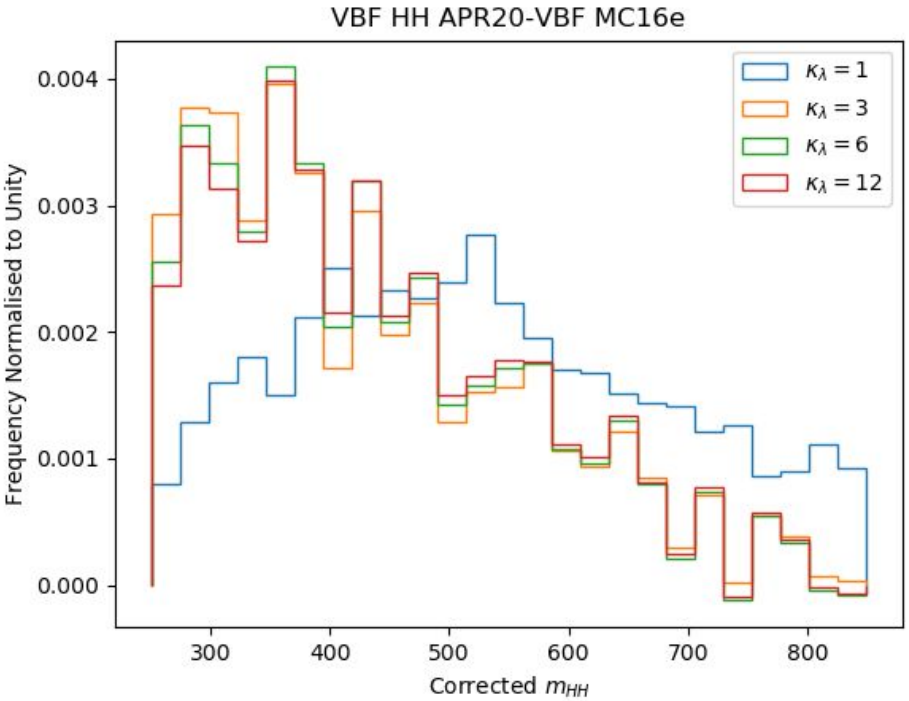
VBF: APR20-VBF. ggF: JUN20. MC16a.
VBF/ggF
$\sigma_{VBFHH}/\sigma_{ggFHH}$
κ_λ
VBF: APR20-VBF. ggF: JUN20. MC16d.
VBF/ggF
$\sigma_{VBFHH}/\sigma_{ggFHH}$
κ_λ
VBF: APR20-VBF. ggF: JUN20. MC16e.
VBF/ggF
$\sigma_{VBFHH}/\sigma_{ggFHH}$
κ_λ

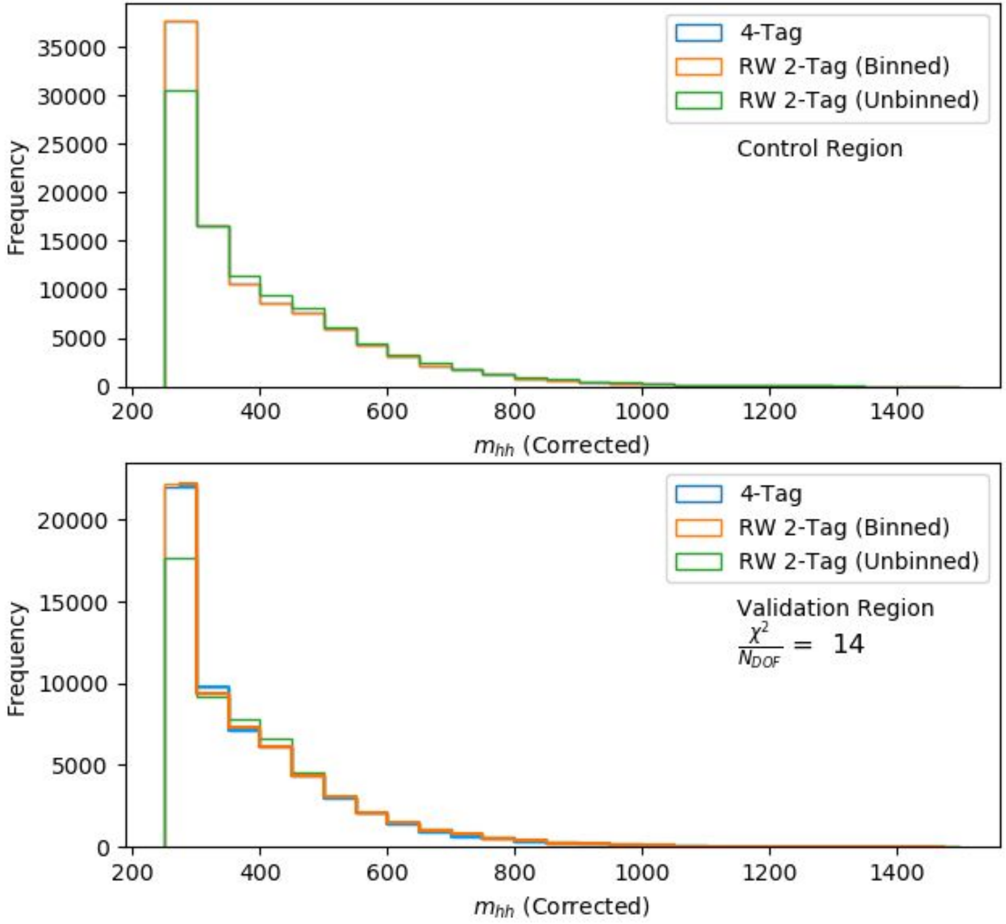

VBF: APR20-VBF. ggF: JUN20. MC16a.
VBF: APR20-VBF. ggF: JUN20. MC16d.
VBF: APR20-VBF. ggF: JUN20. MC16e.
Signal Yield (4-Tag SR)
κ_λ
VBF HH
ggF HH


ggF HH JUN20 MC16e
$\kappa_\lambda = -12$
$\kappa_\lambda = -6$
$\kappa_\lambda = -3$
$\kappa_\lambda = 0$
$\kappa_\lambda = 1$
Frequency Normalised to Unity
0.000
0.001
0.002
0.003
0.004
300
400
500
600
700
800
Corrected m_{HH}


ggF HH JUN20 MC16e
$\kappa_\lambda = 12$
$\kappa_\lambda = 6$
$\kappa_\lambda = 3$
$\kappa_\lambda = 1$
Frequency Normalised to Unity
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
300
400
500
600
700
800
Corrected m_{HH}


VBF HH APR20-VBF MC16e
$\kappa_\lambda = -12$
$\kappa_\lambda = -6$
$\kappa_\lambda = -3$
$\kappa_\lambda = 0$
$\kappa_\lambda = 1$
Frequency Normalised to Unity
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0.0035
300
400
500
600
700
800
Corrected m_{HH}


VBF HH APR20-VBF MC16e
$\kappa_\lambda = 1$
$\kappa_\lambda = 3$
$\kappa_\lambda = 6$
$\kappa_\lambda = 12$
Frequency Normalised to Unity
0.000
0.001
0.002
0.003
0.004
300
400
500
600
700
800
Corrected m_{HH}

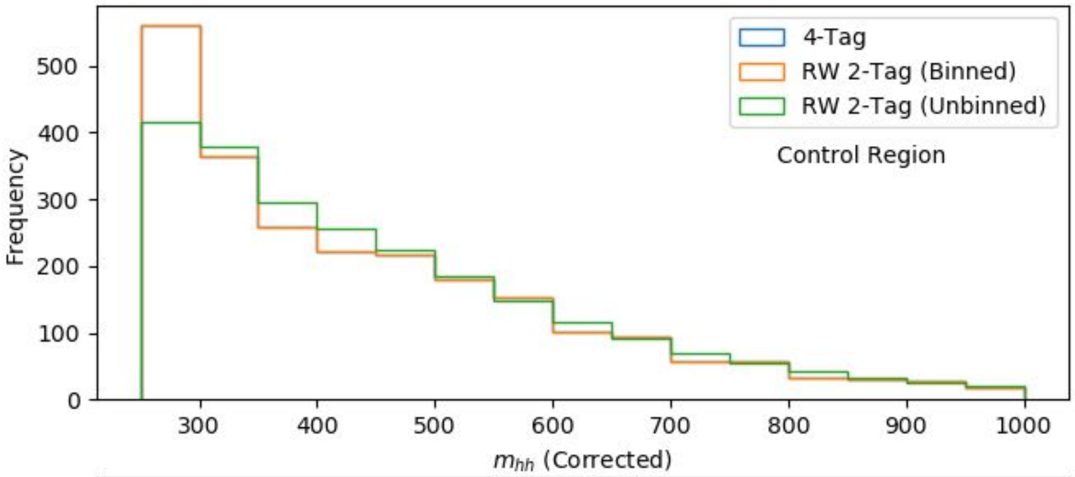


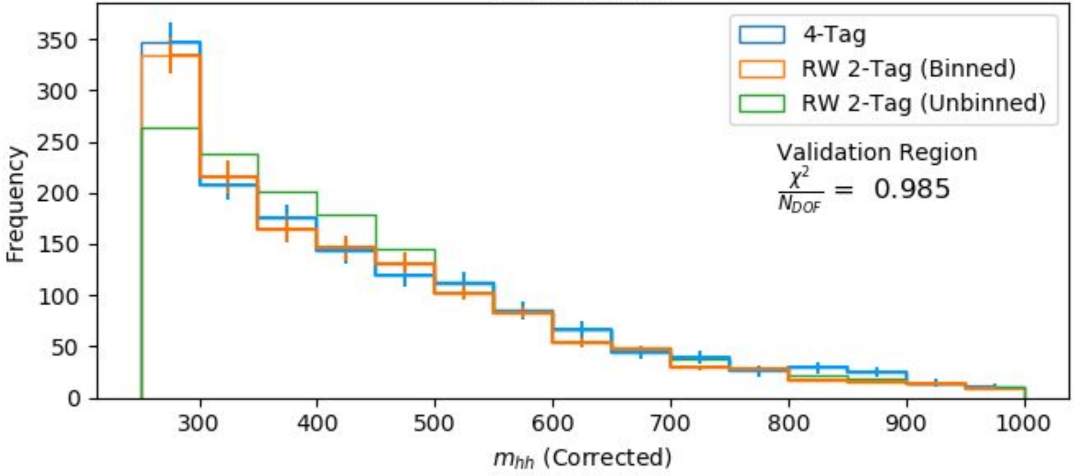
Main (ggF) Selection

JUN20 (2016-2018 Data)

□^{2} =/= 1 → statistical errors are small → motivates the use of an improved method and systematic uncertainty.

High value of □^{2} is driven by single bin at very high mHH bins.





VBF Selection

VBF APR20-VBF (2016-2018 Data)

Range reduced to 0-1000 GeV.

Approximating μ Limit

$$N = \mu S + B \rightarrow \mu = \frac{N-B}{S}$$

Bill's Thesis (6.2.4)

Say signal strength, μ, is a Gaussian r.v. centred at hypothesis μ (as in SM).

For any Gaussian distribution, the below relation holds at the 95% confidence bound :

$$\frac{\mu}{\sigma_\mu} = 1.64$$

Using the top right equation and error propagation, we can generate a formula for the uncertainty on μ, σ_μ, based on the errors on N, B and S.

$$\sigma_\mu = \frac{\sqrt{N}}{S}$$

Knowing σ_μ, the middle equation can be solved numerically to get an approximation of μ at the 95% confidence bound.

Complexity can be increased by including systematics and performing in bins.

VBF and ggF comparison (statistical uncertainty only).

	ggF				VBF			
kl	**S**	**B**	**Binned Sig.**	**Approx. mu**	**S**	**B**	**Binned Sig.**	**Approx. mu**
-6	386.4	43109	2.28	0.72	15.39	1043	0.53	3.17
-5	308.9	43109	1.85	0.89	11.68	1043	0.41	4.17
-4	240.3	43109	1.46	1.13	8.48	1043	0.3	5.71
-3	180.7	43109	1.12	1.47	5.81	1043	0.2	8.29
-2	130.1	43109	0.83	1.98	3.65	1043	0.13	13.06
-1	88.4	43109	0.59	2.79	2.01	1043	0.07	23.25
0	55.8	43109	0.4	4.13	0.89	1043	0.03	50.08
1	32.2	43109	0.25	6.48	0.28	1043	0.01	136.56
2	17.5	43109	0.16	10.58	0.19	1043	0.01	245.29
3	11.9	43109	0.1	16.65	0.61	1043	0.02	83.8
4	15.2	43109	0.09	18.83	1.56	1043	0.05	32.73
5	27.5	43109	0.14	12.17	3.03	1043	0.1	16.79
6	48.8	43109	0.24	6.92	5.02	1043	0.17	10.1
7	79.1	43109	0.39	4.23	7.52	1043	0.25	6.72
8	118.4	43109	0.59	2.8	10.54	1043	0.35	4.78
9	166.7	43109	0.84	1.97	14.08	1043	0.47	3.57
10	224.0	43109	1.13	1.46	18.14	1043	0.61	2.77
11	290.3	43109	1.47	1.12	22.72	1043	0.76	2.21
12	365.5	43109	1.86	0.88	27.82	1043	0.93	1.8
13	449.8	43109	2.3	0.71	33.43	1043	1.12	1.5

VBF and ggF 95% CL limit on μ comparison.

kl	VBF mu limit	ggF mu limit	Ratio
-20	0.52	0.18	0.34
-19	0.57	0.19	0.34
-18	0.63	0.21	0.34
-17	0.7	0.24	0.34
-16	0.78	0.26	0.34
-15	0.87	0.29	0.33
-14	0.99	0.33	0.33
-13	1.13	0.37	0.32
-12	1.3	0.42	0.32
-11	1.51	0.48	0.32
-8	2.6	0.76	0.29
-7	3.22	0.92	0.28
-6	4.12	1.12	0.27
-5	5.43	1.39	0.26
-4	7.5	1.78	0.24
-2	17.41	3.2	0.18
-1	31.52	4.7	0.15
0	69.39	7.22	0.1
1	161.37	11.88	0.07

kl	VBF mu limit	ggF mu limit	Ratio
2	241.88	18.88	0.08
3	90.4	25.83	0.29
4	37.12	25.5	0.69
5	19.58	16.75	0.86
6	11.97	9.67	0.81
7	8.05	5.87	0.73
8	5.77	3.9	0.68
9	4.34	2.73	0.63
10	3.38	2.03	0.6
11	2.7	1.56	0.58
12	2.21	1.23	0.56
13	1.84	1	0.54
14	1.56	0.82	0.53
15	1.34	0.69	0.52
16	1.16	0.59	0.51
17	1.01	0.51	0.5
18	0.89	0.44	0.5
19	0.79	0.39	0.49
20	0.71	0.34	0.48

NB: points -10, -9, -3 are missing.