HH→bbττ Analysis Overview

Two final states

- $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$

VBF-specific baseline selection

- Require 2 VBF jets in the event (on top of ggF selection)

The VBF analysis updated with the latest CxAOD production

- RNN loose τ ID, EMPFlow jets, DL1r tagging with 77% WP

Trigger dependent $\tau_{lep}\tau_{had}$ ggF selection

- Single lepton (e or μ) trigger (SLT)
  - $p_{T\,lepton} > 25\text{-}27$ GeV, $p_T^{\tau} > 20$ GeV
- SR has 1 **τ**, 1 e or 1 μ + 2 central b-tagged jets, $m_{bb} < 150$ GeV and $m_{\tau\tau}^{MMC} > 60$ GeV
- In the process of adding lepton τ trigger channel (LTT)

Trigger dependent $\tau_{had}\tau_{had}$ ggF selection

- Single **τ** trigger (STT)
  - $p_T^{\tau 1} > 100\text{-}180$ GeV, $p_T^{\tau 2} > 20$ GeV
- Di-**τ** trigger (DTT)
  - For 2015 and 2016:
    - $p_T^{lead} > 80$ GeV, $p_T^{sublead} > 20$ GeV
  - For 2017 and 2018:
    - L1Topo: $p_T^{sublead} < 45$ GeV and $\Delta R_{\tau\tau} < 2.5$
    - 4J12: $p_T^{sublead} > 45$ GeV
- SR has 2 **τ** + 2 central b-tagged jets, and $m_{\tau\tau}^{MMC} > 60$ GeV

L1Topo:
HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1DRTAU20ITAU12I-J25
4j12:
HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM_4J12.0ETA23

Jet Selection

Forward Jets:

- $p_T > 30$ GeV
- $|\eta| > 2.5$

Central Jets:

- $p_T > 20$ GeV
- $|\eta| < 2.5$

All Jets

VBF Jets:

Take highest m_{jj} pair with $\eta_{j1}{}^*\eta_{j2} < 0$ (non b-tagged)

Remove selected VBF Jets from the jet bucket

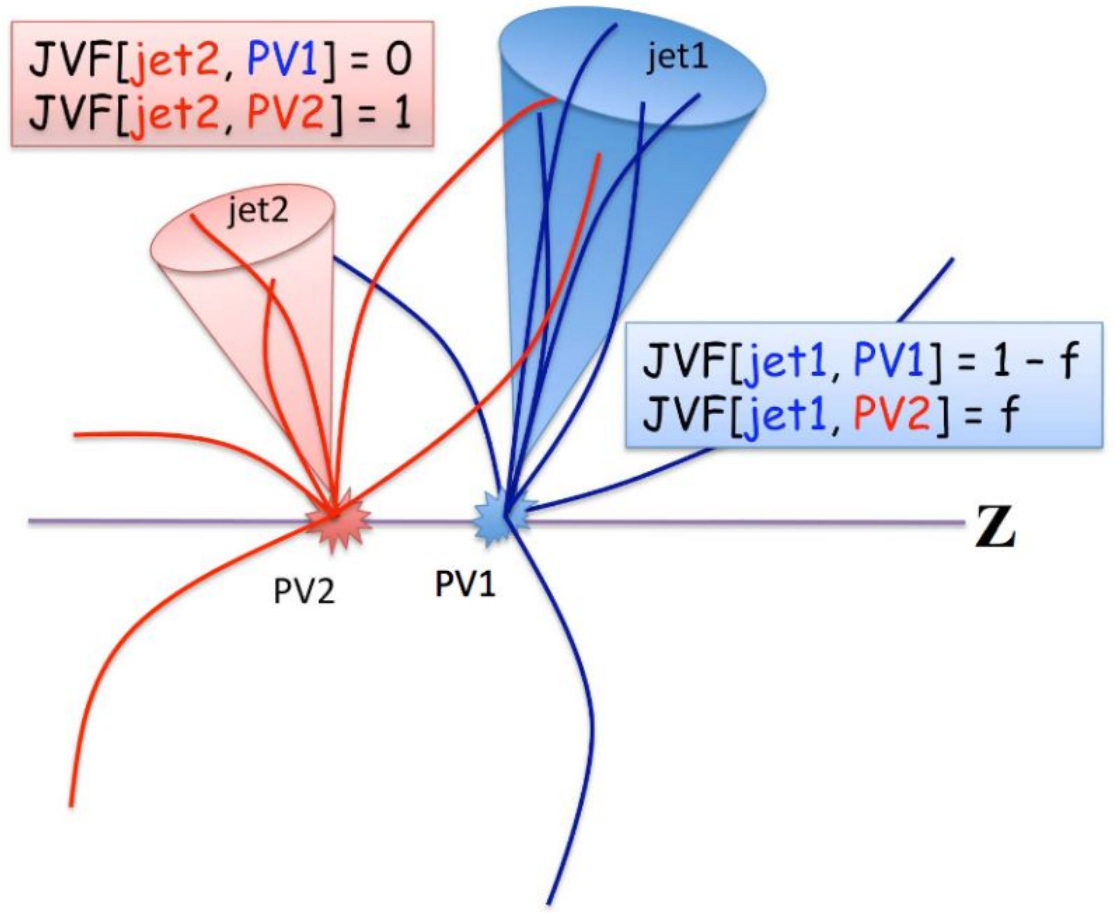
Select Higgs jets from the remaining central jets

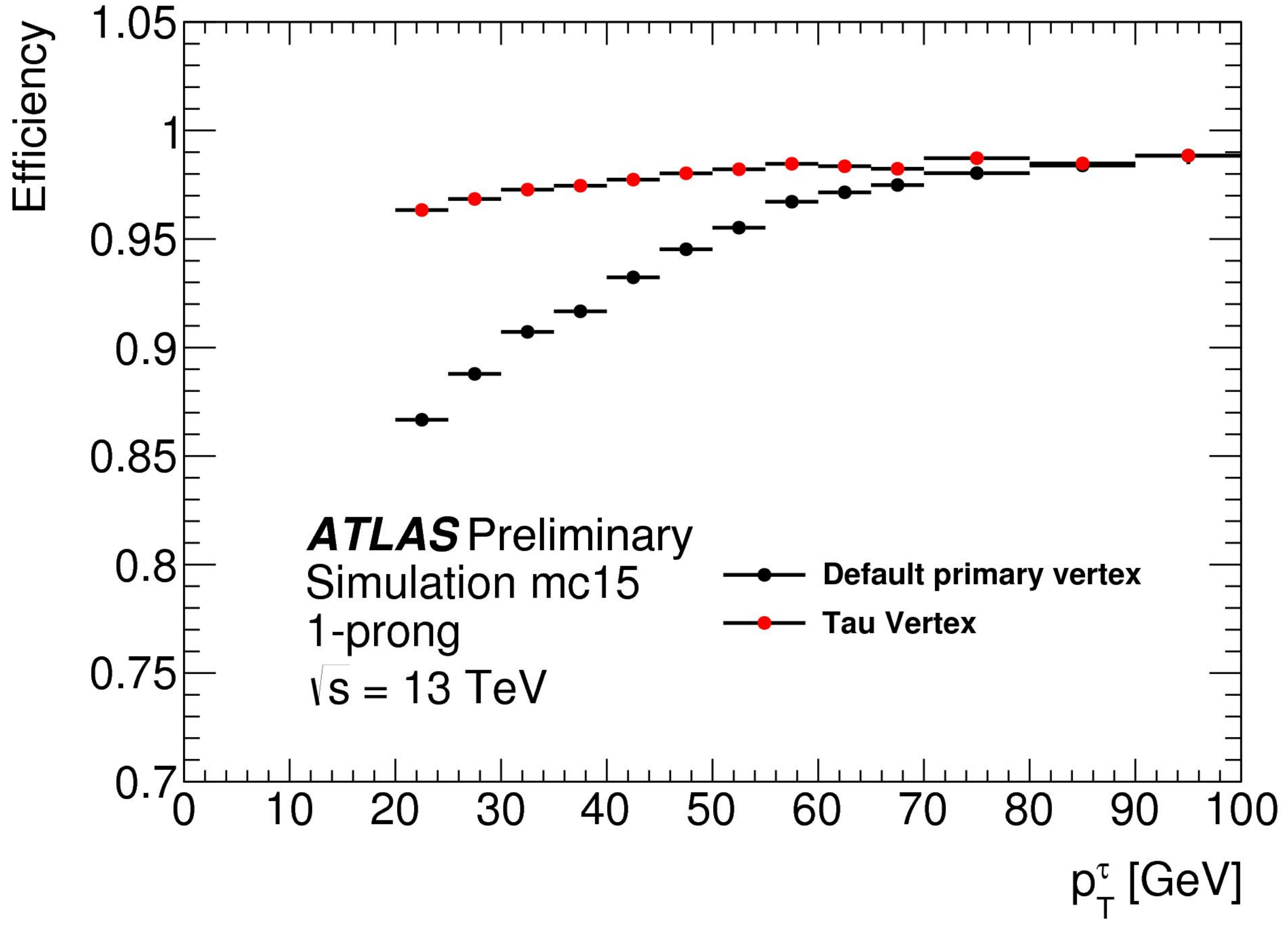
τhadτhad: Tau Vertex Requirement

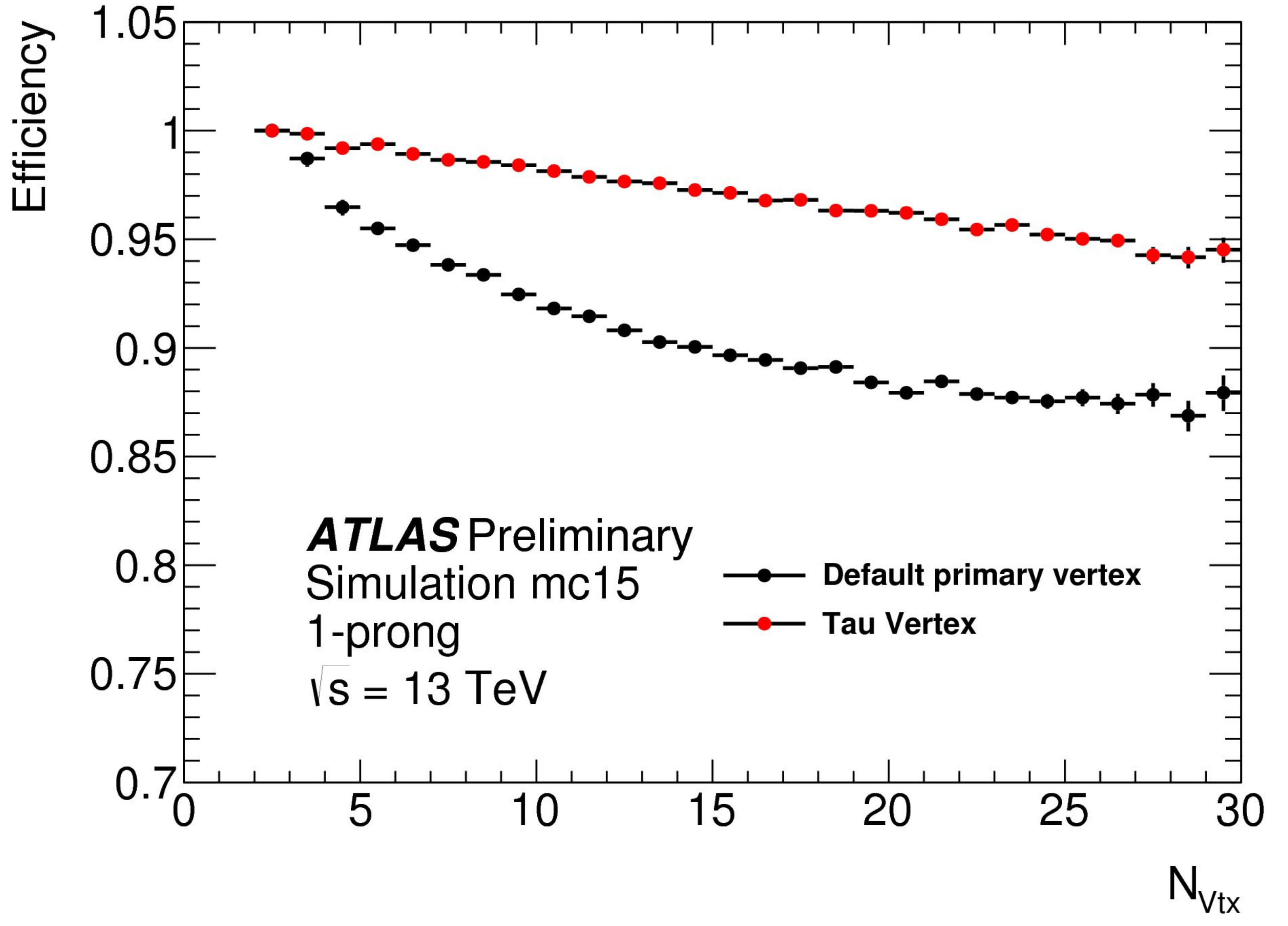
- Tau-jets have their own vertex association algorithm compared to other ATLAS objects

- The vertex with the highest tau-JVF rather than the hard scatter (HS) vertex is taken

$$\mathrm{TJVF}(\tau_i, \mathrm{vtx}_j) = \frac{\sum_k p_T(\mathrm{trk}_k^{\tau_i}, \mathrm{vtx}_j)}{\sum_{n,l} p_T(\mathrm{trk}_l^{\tau_i}, \mathrm{vtx}_n)}.$$







- Although it won't happen very often with our selection, there will be events where the b-jets are not associated to the same vertex as the taus

ThadThad: Yields Comparison

Apply vertex requirement that vertex match for all objects

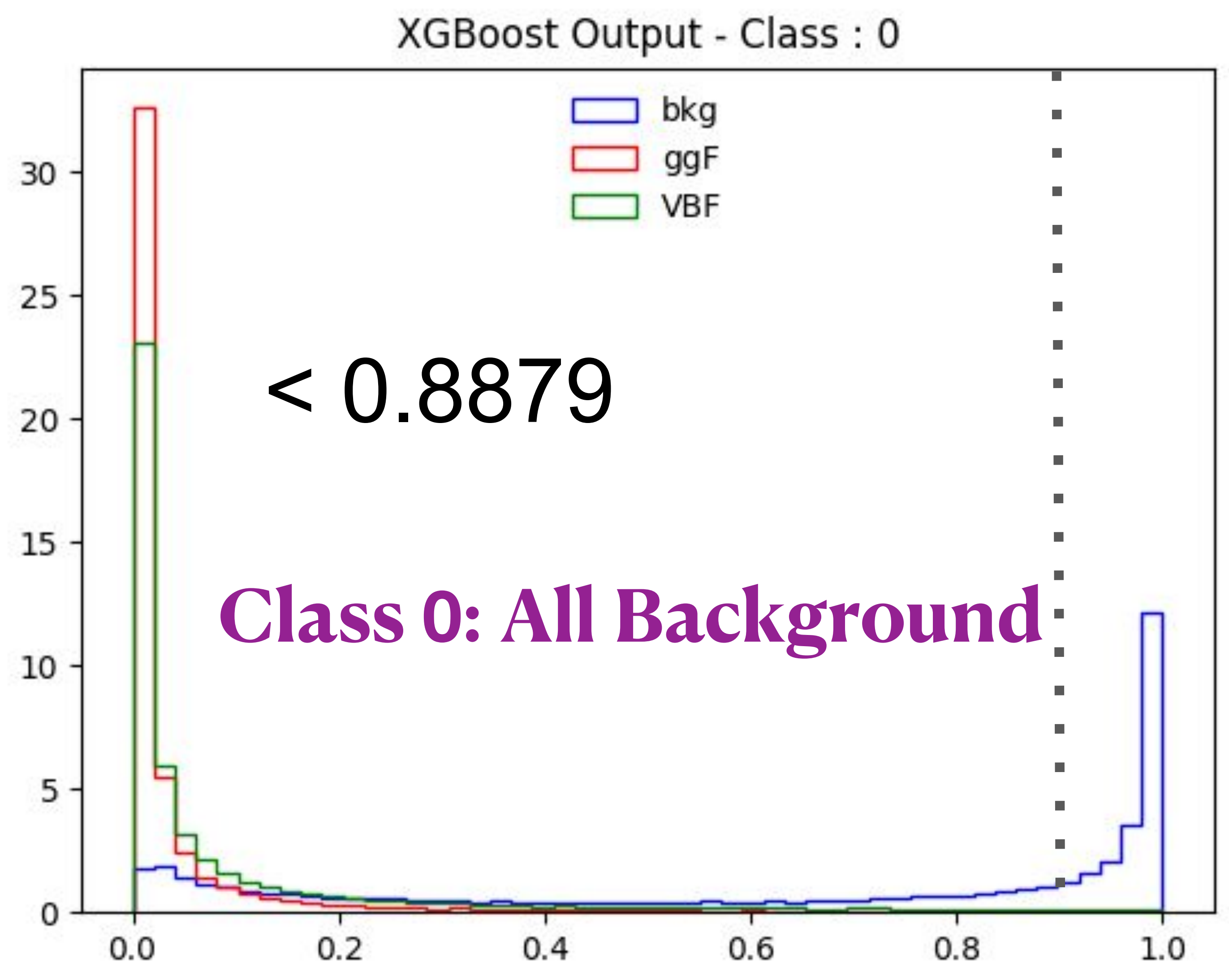
no vtx	CR yield	SR yield	$s/\sqrt{bkg}$
SM signal	0.0216105	0.107536	0.001834
$c_{2V} = 0$	0.503521	3.53399	0.060262
$c_{2V} = 0.5$	0.157788	1.08654	0.018528
$c_{2V} = 1.5$	0.0856676	0.576275	0.009827
$c_{2V} = 2$	0.353655	2.48101	0.042307
$c_{2V} = 4$	3.51443	24.5828	0.419192
bkg	41599.4	3439.03	
data	41021	2897	

With vertex requirement

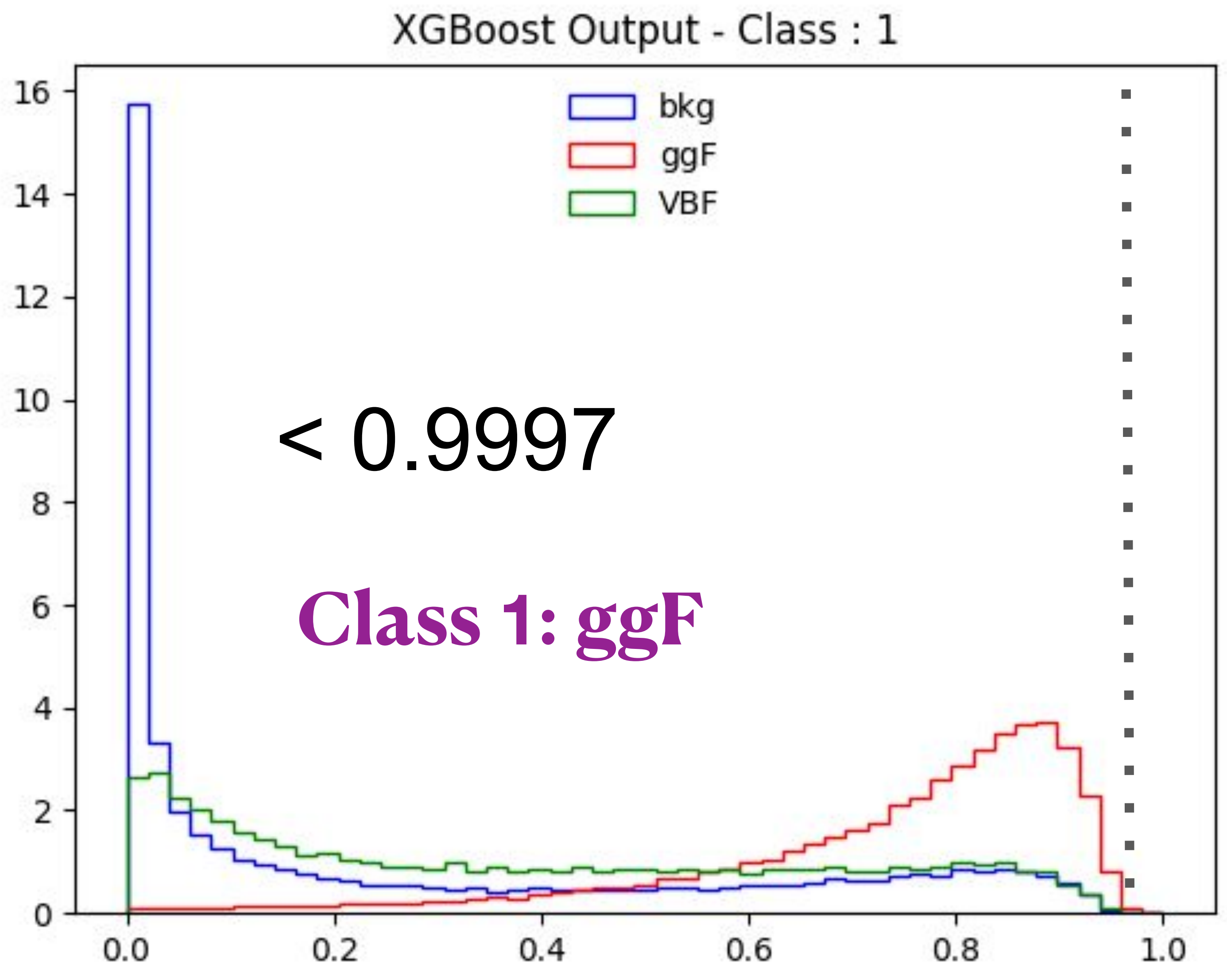
with vtx	CR yield	SR yield	$s/\sqrt{bkg}$
SM signal	0.0214477	0.107332	0.001878
$c_{2V} = 0$	0.49864	3.52559	0.061687
$c_{2V} = 0.5$	0.156649	1.08354	0.018959
$c_{2V} = 1.5$	0.0849659	0.574834	0.010058
$c_{2V} = 2$	0.351149	2.47637	0.043329
$c_{2V} = 4$	3.48237	24.5344	0.429281
bkg	36671.7	3266.43	
data	36175	2775	

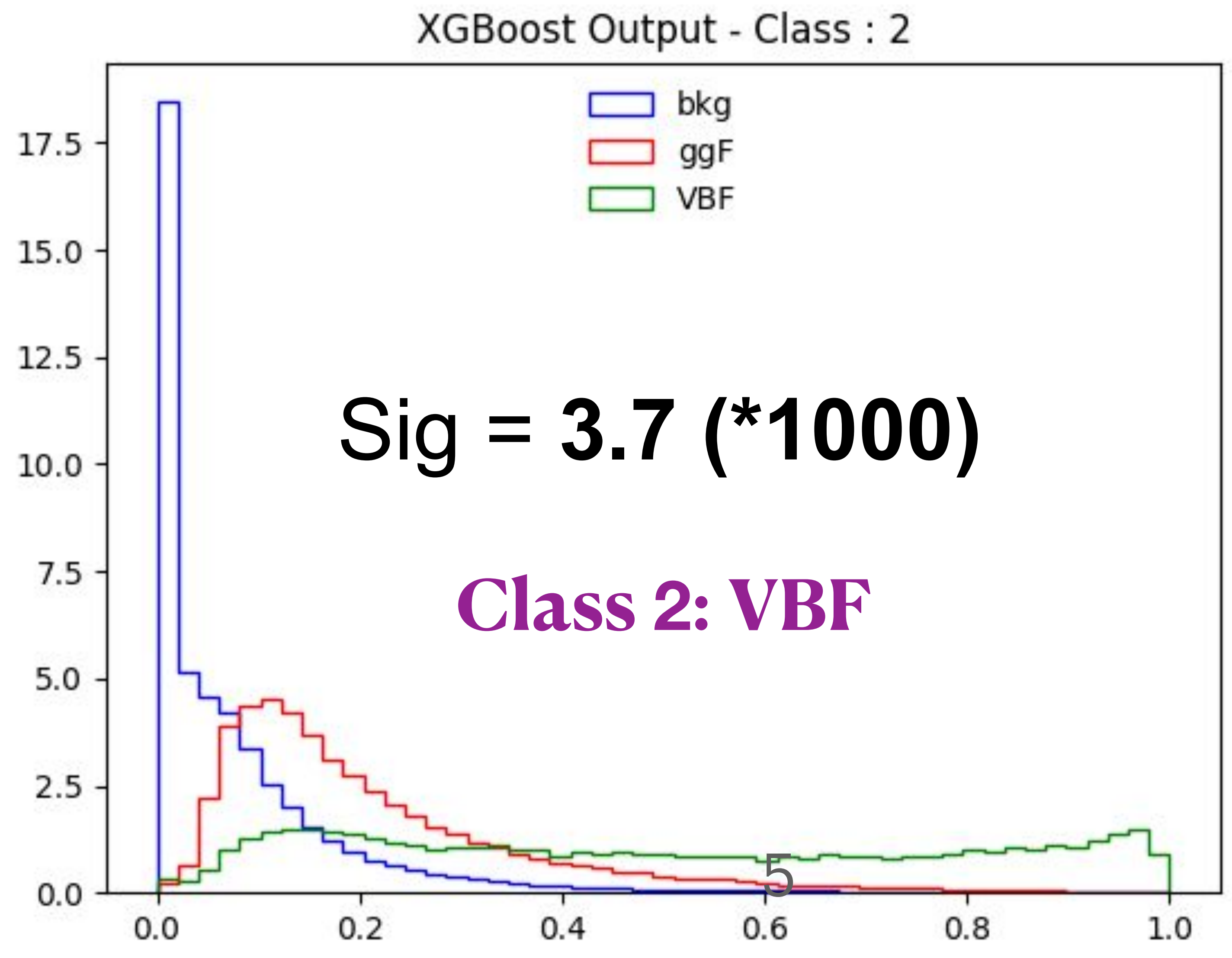
Impact of tau vertex requirement: better background rejection, improving significance by ~2-4 %

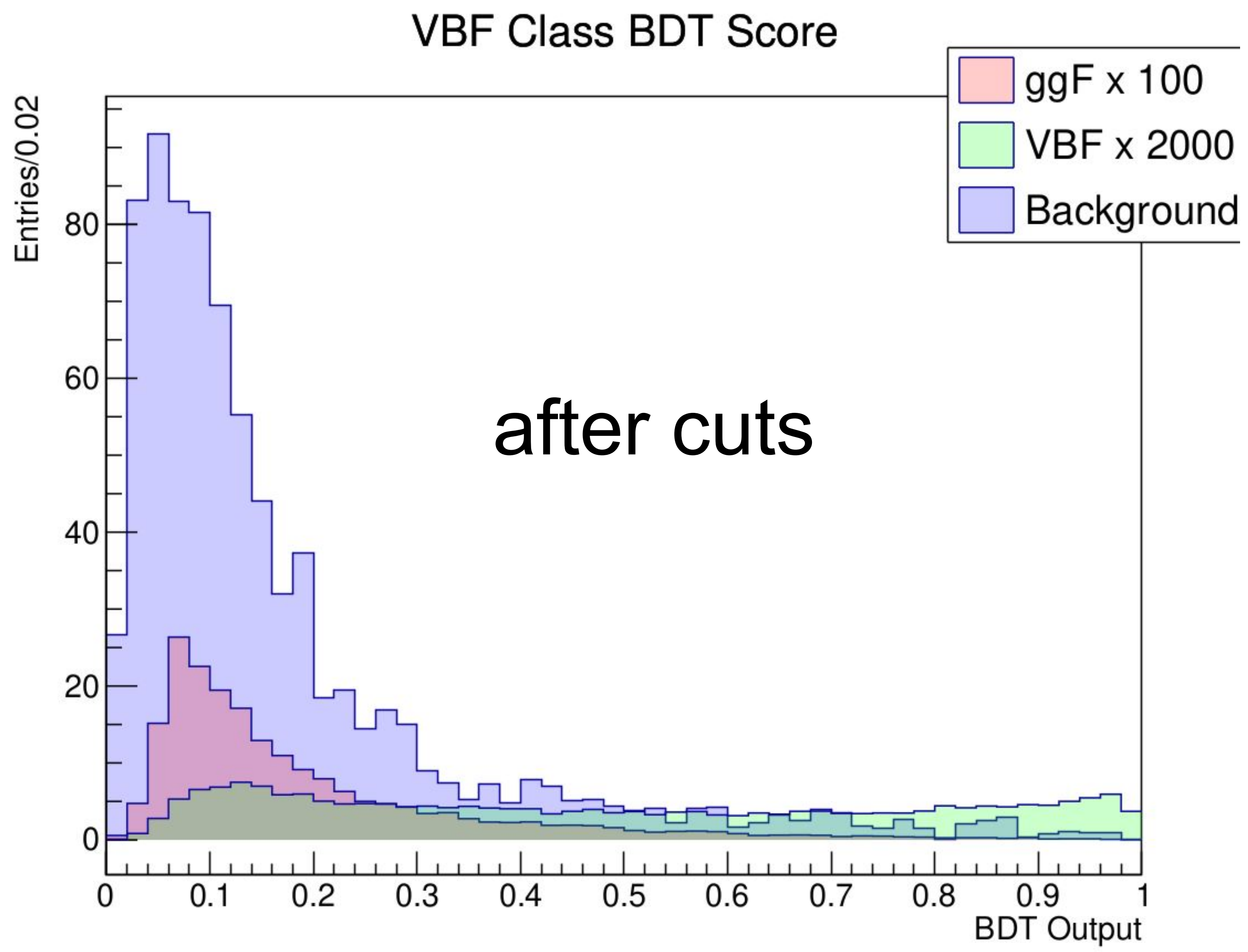
$\tau_{had}\tau_{had}$: Multiclass BDT Output (Very Preliminary)



- First pass at training a multiclass BDT (using xgboost) using a similar strategy to $bb\gamma\gamma$
- Total of 27 variables : ggF analysis + event shape variables + VBF variables
  - These will be reduced and a smaller set will be used in the future - highly correlated
- Strategy to cut on background score (< 0.005) & ggF score (< 0.65) to maximise significance to use VBF score as a final discriminant
  - Gives the significance of ~50 (*1000) compared to ~2.4 (*1000) with grid search
    - However this leaves very little in the SR, 0.00003% of the background remains…
  - Requiring a minimum of 20% Bkg remaining was implemented before optimising just for checks. Results shown in the plots







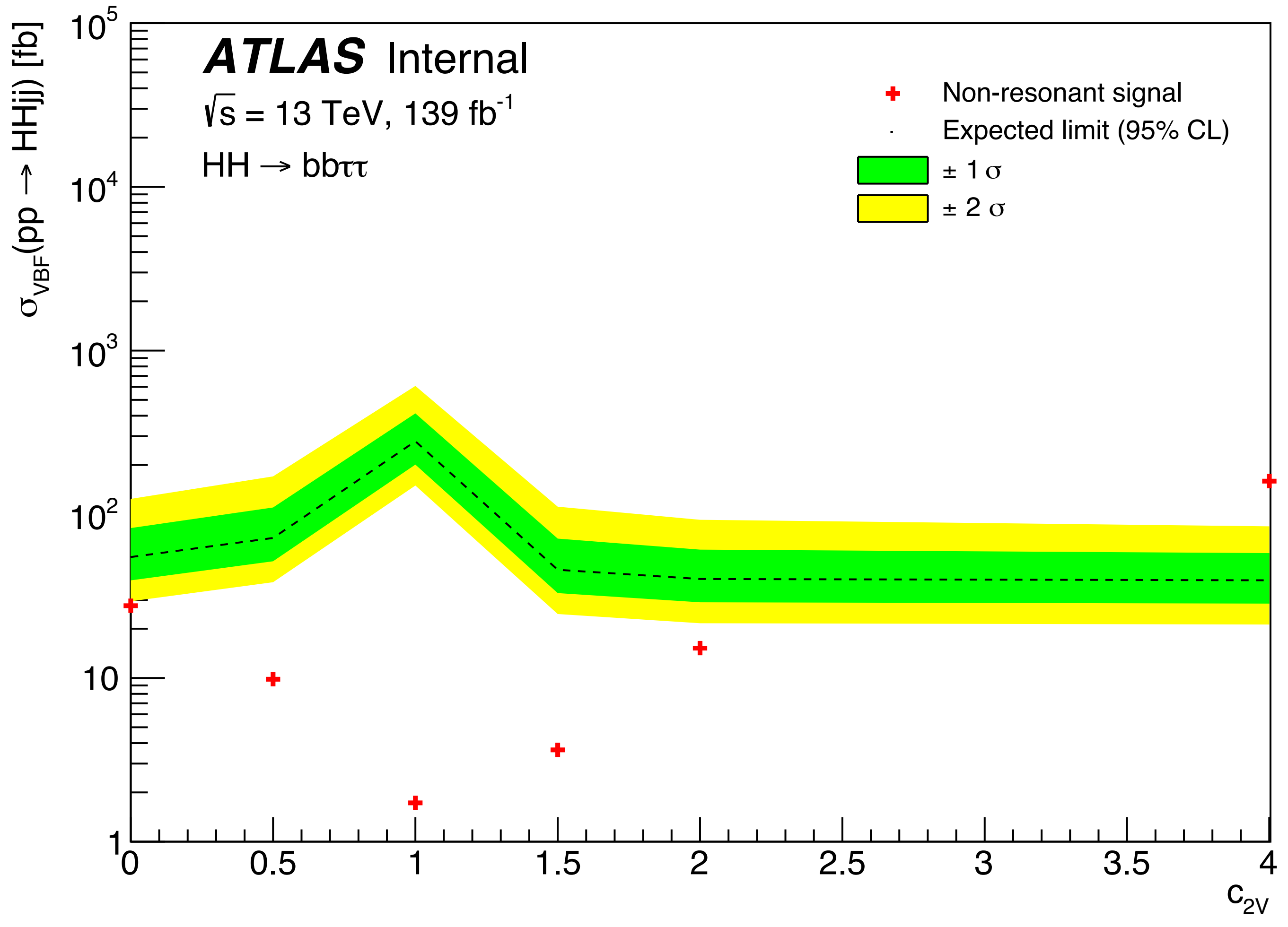
$\tau_{had}\tau_{had}$: BDT Studies for c_{2V} Scan

Dorothee Dapper

- BDT studies performed (independently from previously shown studies) to set preliminary limits on c_{2V}
- VHH contamination is taken into account
- Input variables: m_{HH}, m_{jj}, $\Delta\eta_{jj}$, m_{bb}, m_{MMC} $X_{m_{bb}m_{\tau\tau}} = \sqrt{\left(\frac{m_{bb}-112\text{ GeV}}{0.1\cdot m_{bb}}\right)^2 + \left(\frac{m_{\tau\tau}-116\text{ GeV}}{0.1\cdot m_{\tau\tau}}\right)^2}$ $X_{\Delta R_{bb}\Delta R_{\tau\tau}} = \sqrt{\left(\frac{\Delta R_{bb}-1.227}{0.1\cdot\Delta R_{bb}}\right)^2 + \left(\frac{\Delta R_{\tau\tau}-1.229}{0.1\cdot\Delta R_{\tau\tau}}\right)^2}$

Limits on the VBF HH production cross section $\sigma_{\mathrm{VBF}}(pp \to HHjj)$
dependent on the VVHH coupling c_{2V}

c_{2V}	μ_{VBF}	$\sigma_{\mathrm{VBF}}(pp \to HHjj)$
0.0	2.0	55 fb
0.5	7.3	72 fb
1.0	161.7	279 fb
1.5	12.6	46 fb
2.0	2.6	40 fb
4.0	0.2	40 fb



Summary & Next Steps

Updates to the analysis presented

- Include tau vertex requirement
- Multiclass BDT implemented with xgboost similar to bbyy
  - Good performance on separation seen
  - Cutting on Bkg and ggF BDT scores gives a conservative significance increase from 2.4 (*1000) → 3.7 (*1000) compared to the grid search

Limits on c_{2v}

- Checks done in hadhad channel, not able to exclude c_{2v}=0 but close
- The studies are ongoing for the estimation of the limits on c_{2v}=0 in lephad channel

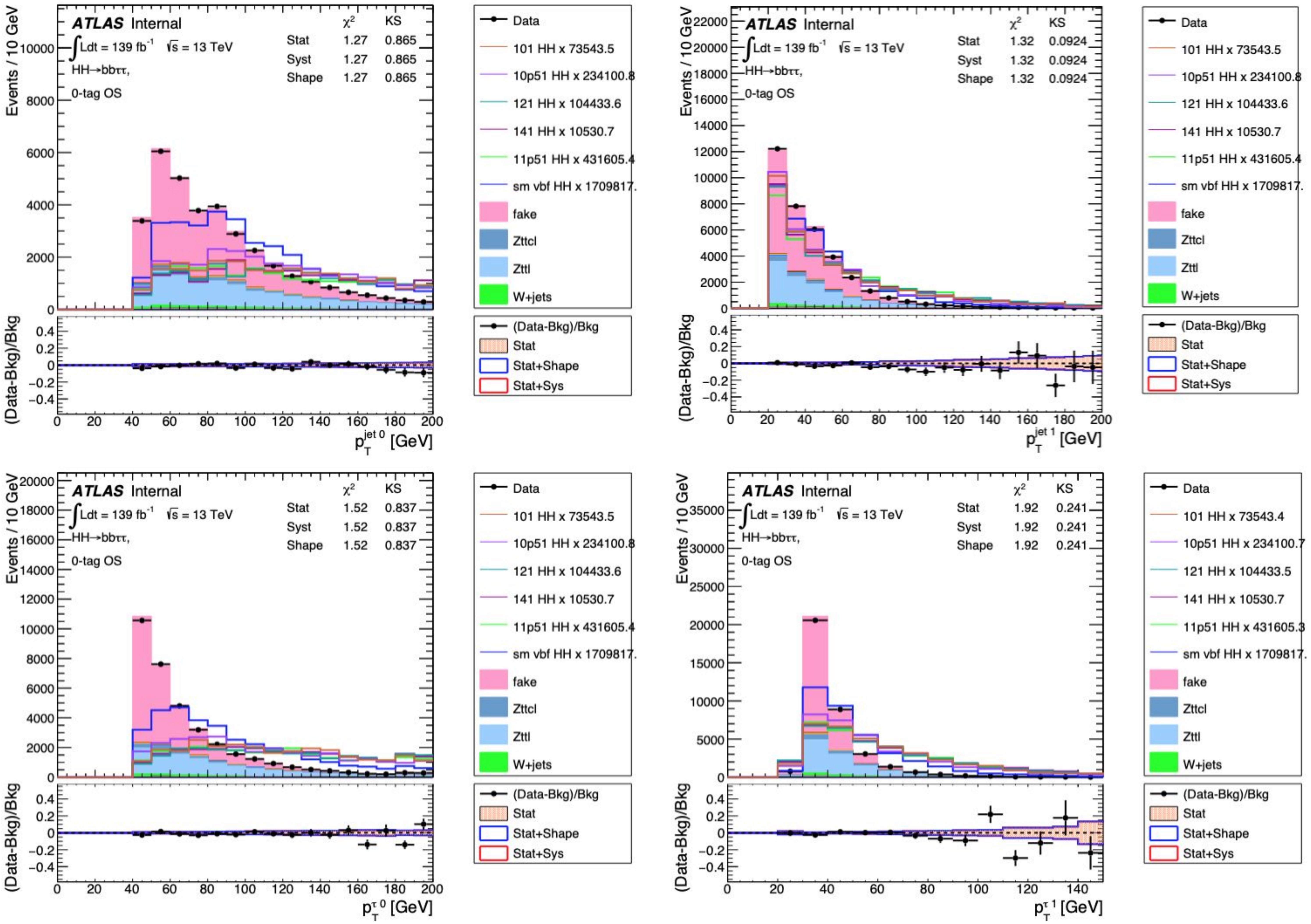
Next steps

- Study the orthogonality between ggF and VBF categories
- Studies on c_{2v} scans are still ongoing
- Update the analysis with the latest production with systematics included

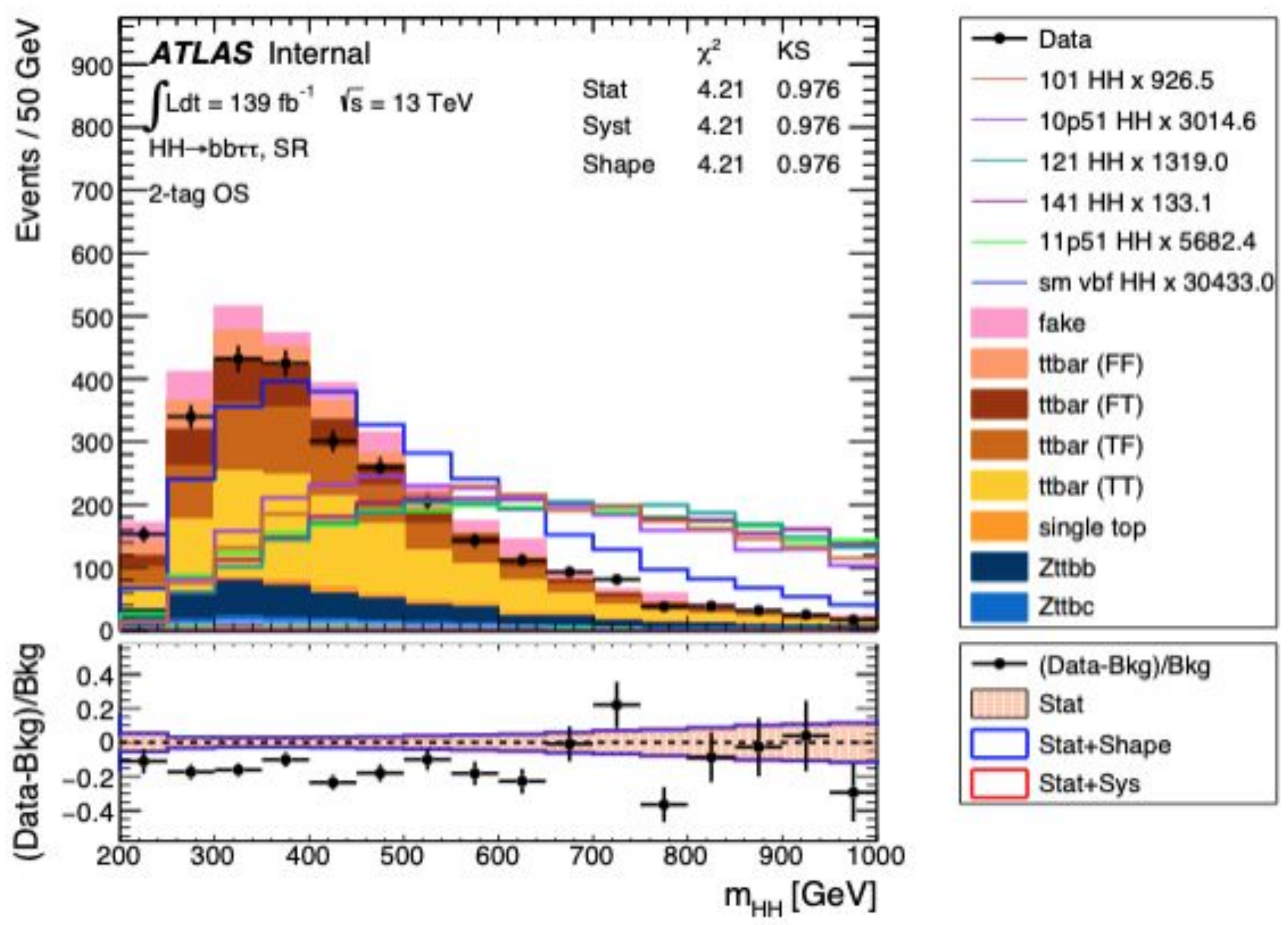
BACK UP

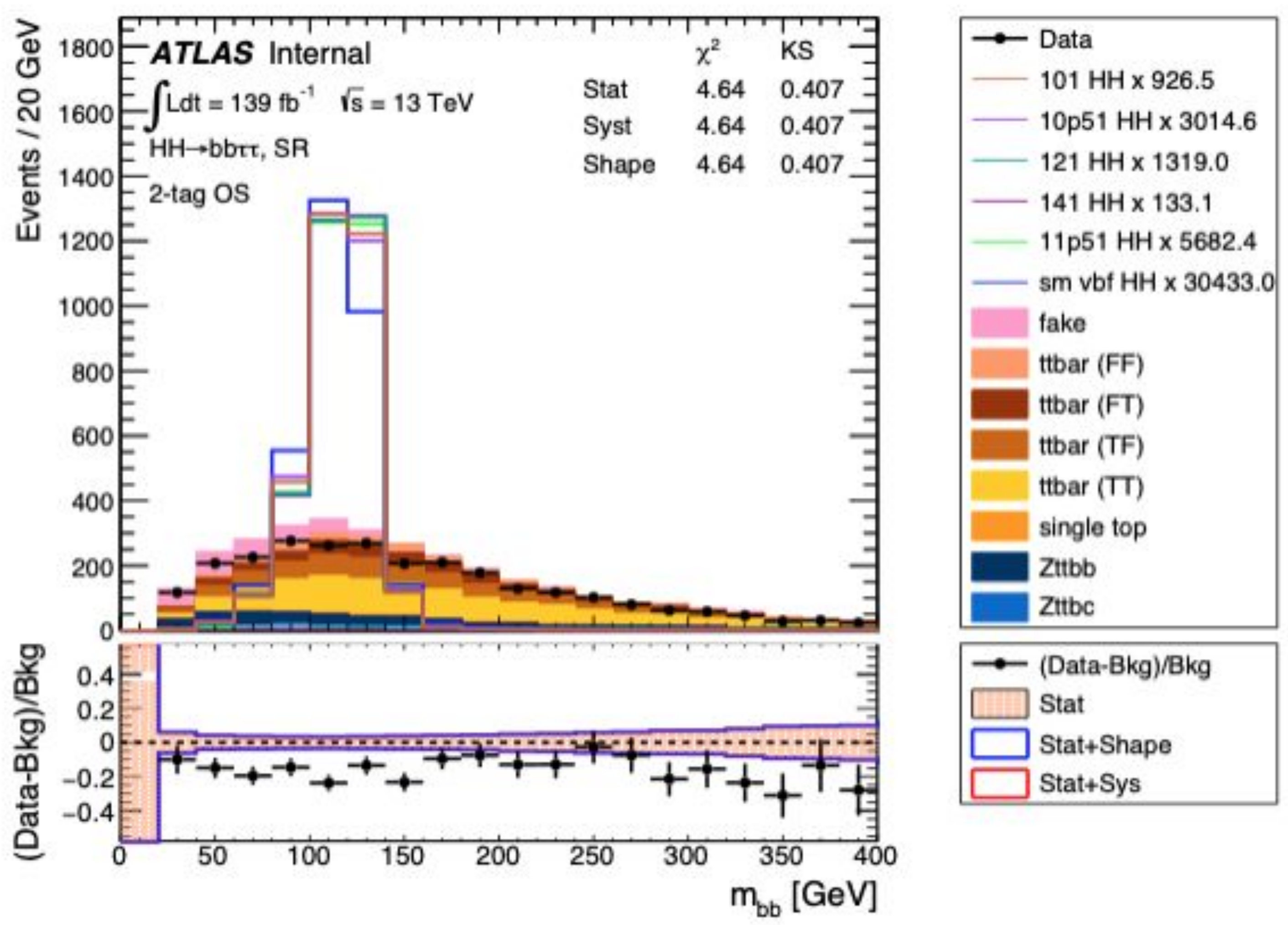
$\underline{\tau_{had}\tau_{had}}$: Leading, Subleading Jet and Tau p_T @ 0-Tag OS Region

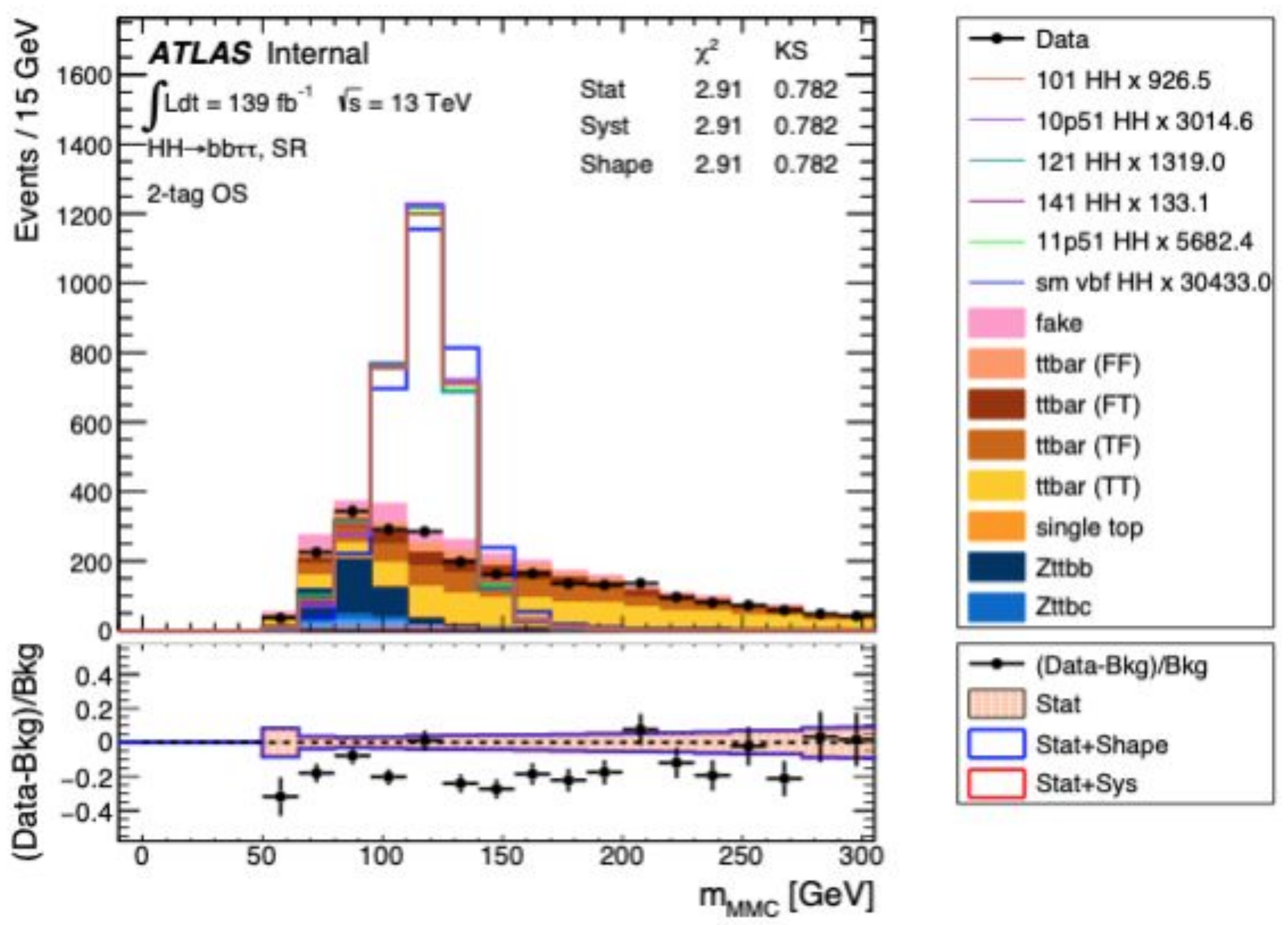
- VBF SM and c_{2v} varied couplings signals along with main backgrounds
- acceptable closure in the CR

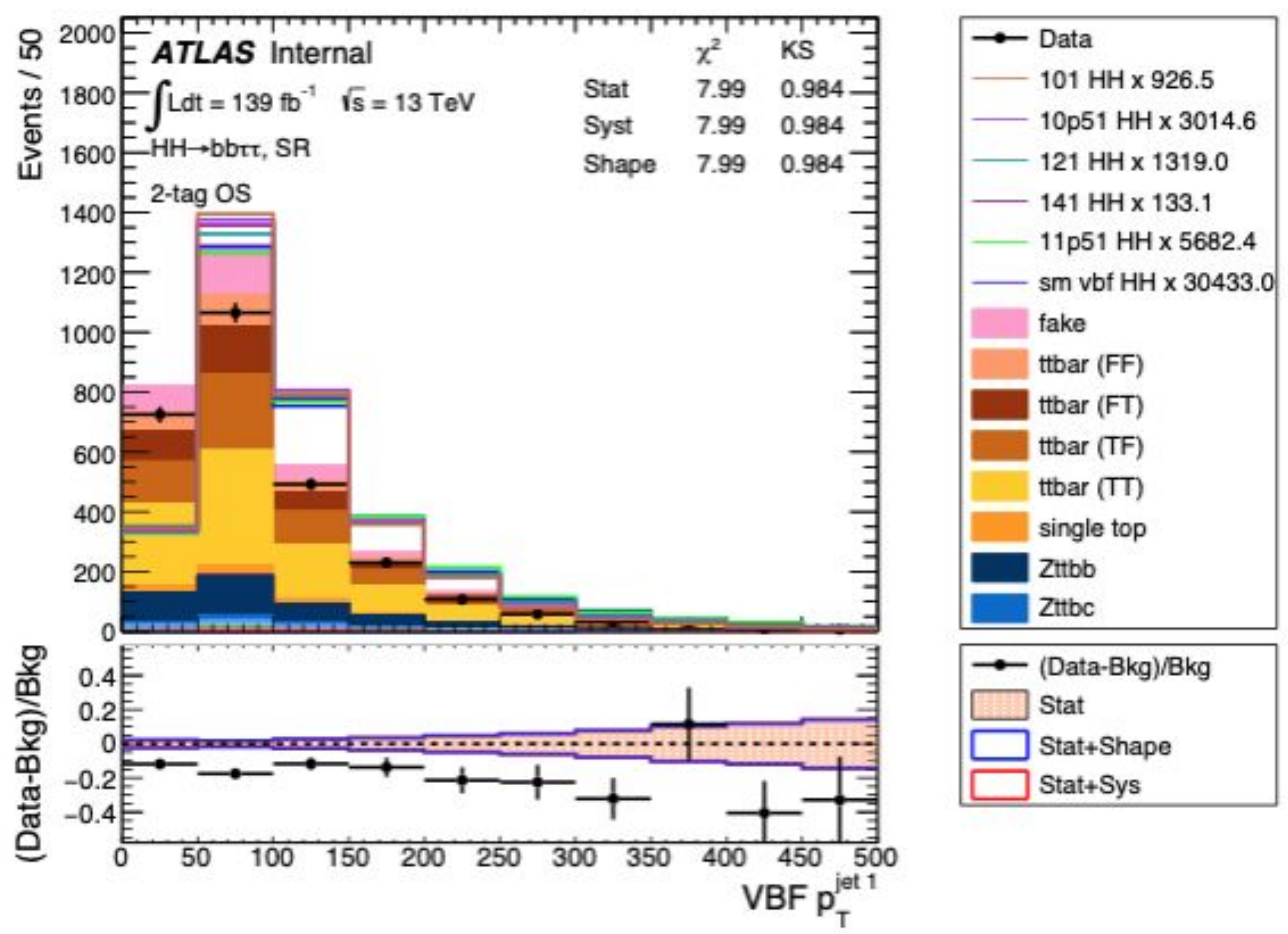


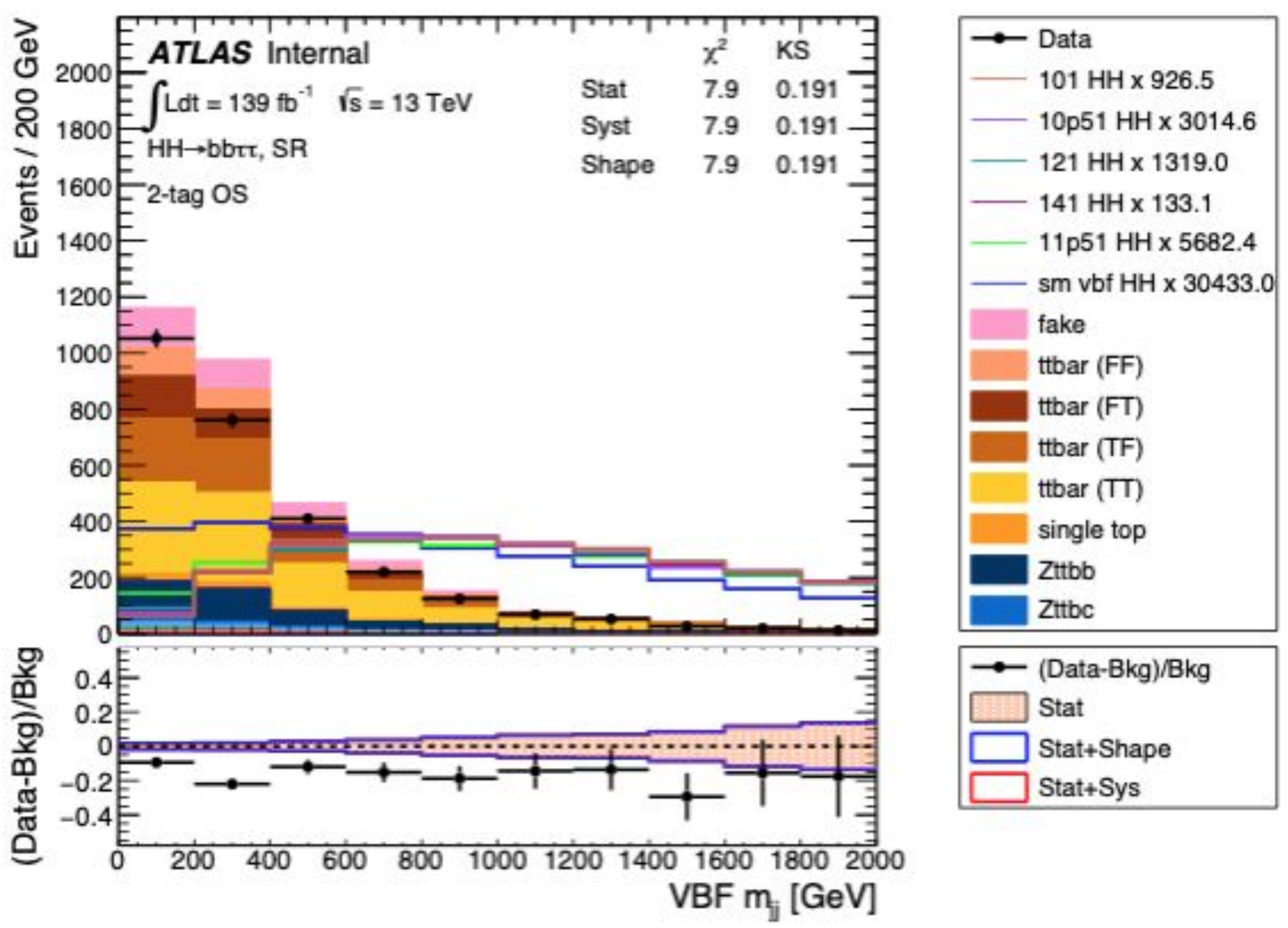
$\tau_{had}\tau_{had}$: Relevant Distributions in the SR

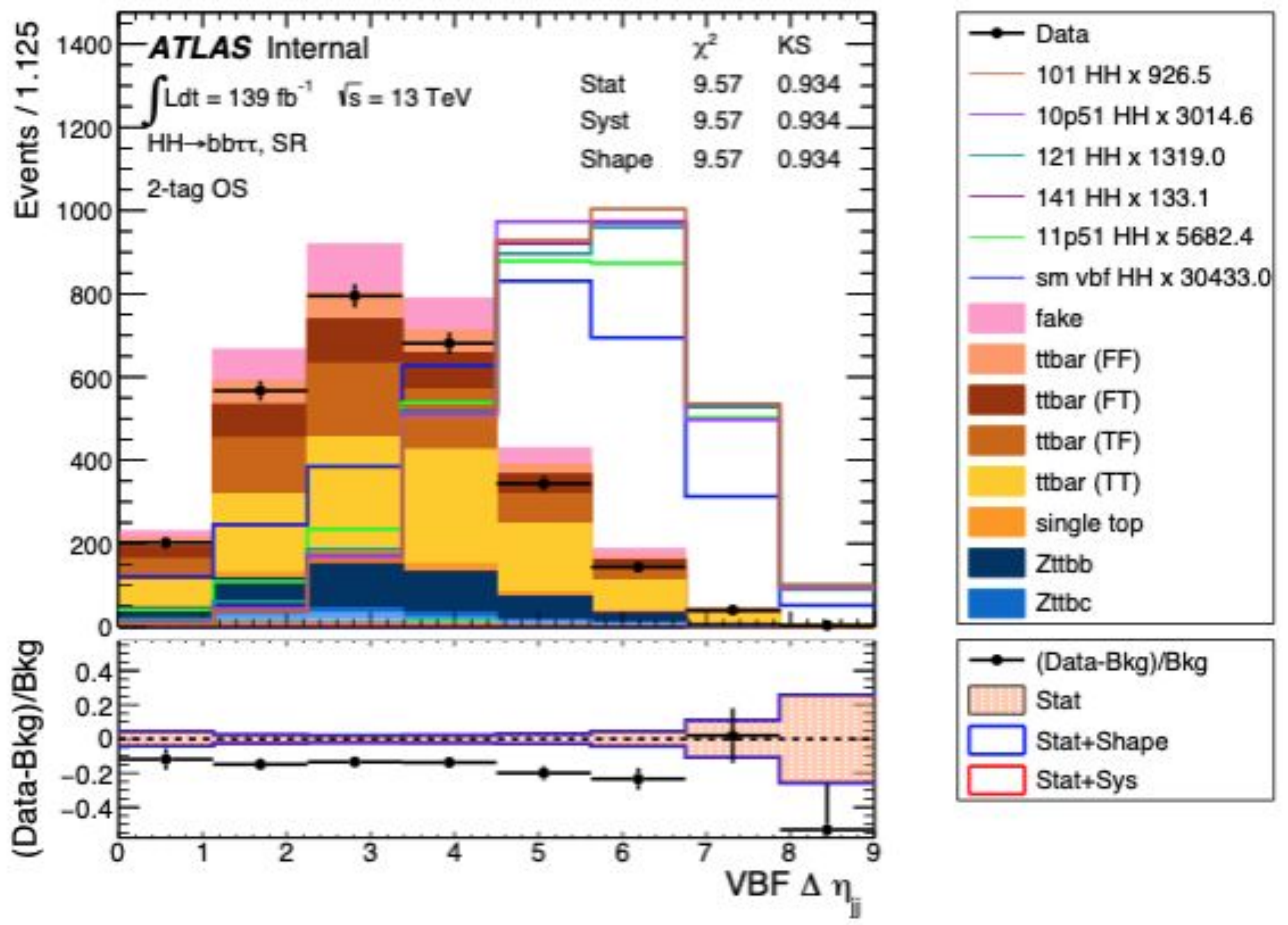












ThadThad: Grid Optimisation Studies

- Evaluating a set of exclusive VBF tagging jets topology cuts: M_{jj} , $\Delta\eta_{jj}$ and p_T^{j1} :

  mjjcuts = ["300", "350", "400", "450","500", “550”,"600","650","700","800","1000","1200"]

  detajjcuts = [“2.5”,”3.0”, “3.5”,”4.0”,"4.5","5.0"]

  $p_Tj_1$cuts = ["30","35","40","45","50", “55”,”60","65","70","75","80","85","90","95","100"]

- Applying a yield threshold of 0.08 on the VBF SM signal (choice based on maximum yield obtained with previous tagger wp), see next slide

$\tau_{had}\tau_{had}$: Optimisation of SM VBF Signal

With vertex requirement

M_{jj}, $\Delta\eta_{jj}$ and p_T^{j1}

M_{jj} [GeV]	$\Delta\eta_{jj}$	$p_T^{j_1}$ [GeV]	signal yield	bkg yield	$(s/\sqrt{b}) \times 1000$
300	3	30	0.0847	1569.3952	2.1382
300	3	35	0.0839	1514.0492	2.1558
300	3	40	0.0824	1451.9077	2.1616
300	3	45	0.0802	1385.1504	2.1554
300	3.5	30	0.0799	1357.5991	2.1689
350	3	30	0.0826	1386.4864	2.2188
350	3	35	0.0819	1338.6577	2.2393
350	3	40	0.0805	1286.7499	2.2454
400	3	30	0.0801	1218.2420	2.2956

M_{jj} [GeV]	$\Delta\eta_{jj}$	$p_T^{j_1}$ [GeV]	signal yield	bkg yield	$(s/\sqrt{b}) \times 1000$
300	3	30	0.0844	1463.5535	2.2074
300	3	35	0.0836	1415.6622	2.2225
300	3	40	0.0821	1358.8130	2.2275
300	3	45	0.0800	1297.6690	2.2198
300	3.5	30	0.0797	1268.2020	2.2374
350	3	30	0.0824	1295.4151	2.2882
350	3	35	0.0817	1253.9362	2.3063
350	3	40	0.0803	1206.6916	2.3113
400	3	30	0.0799	1138.6439	2.3670

M_{jj} and $\Delta\eta_{jj}$

M_{jj} [GeV]	$\Delta\eta_{jj}$	signal yield	bkg yield	$(s/\sqrt{b}) \times 1000$
300	2.5	0.0874	1688.8440	2.1277
300	3	0.0847	1569.1968	2.1384
300	3.5	0.0799	1357.4006	2.1691
350	2.5	0.0849	1471.0547	2.2135
350	3	0.0826	1386.4864	2.2188
400	2.5	0.0820	1278.0288	2.2938
400	3	0.0801	1218.2420	2.2956
450	2.5	0.0790	1124.4253	2.3567
450	3	0.0774	1080.1048	2.3566

M_{jj} [GeV]	$\Delta\eta_{jj}$	signal yield	bkg yield	$(s/\sqrt{b}) \times 1000$
300	2.5	0.0872	1576.6125	2.1954
300	3	0.0844	1463.3448	2.2075
300	3.5	0.0797	1267.9933	2.2375
350	2.5	0.0846	1376.7126	2.2808
350	3	0.0824	1295.4151	2.2882
400	2.5	0.0817	1196.7411	2.3630
400	3	0.0799	1138.6439	2.3670
450	2.5	0.0788	1053.3479	2.4276
450	3	0.0772	1010.5050	2.4289

Overall significance improvement when adding VBF tagging jets cuts to the baseline selection including tau vertex requirement (showing here only main backgrounds: Ztt, ttbar and fakes)

$\tau_{had}\tau_{had}$: Optimisation of c_{2V} Varied Coupling Signals

- Applying a different yield threshold for each c_{2V} coupling value, (choice based on maximum yield obtained with previous tagger wp)

70% wp	CR yield	SR yield	$s/\sqrt{bkg}$
SM signal	0.0298069	0.0848596	0.001689
$c_{2V}=0$	0.757824	2.84747	0.056674
$c_{2V}=0.5$	0.237998	0.869876	0.017313
$c_{2V}=1.5$	0.130142	0.457537	0.009106
$c_{2V}=2$	0.540383	1.98646	0.039537
$c_{2V}=4$	5.3018	19.7088	0.392270
bkg	39320.8	2524.35	
data	39320.8	2113	

M_{jj} [GeV]	$\Delta\eta_{jj}$	$p_T^{j_1}$ [GeV]	signal yield	bkg yield	$(s/\sqrt{b})$
600	3.5	35	$(c_{2V}=0.0)$2.8693	679.0698	0.1101
600	3.5	35	$(c_{2V}=0.5)$0.8709	679.0698	0.0334
600	2.5	30	$(c_{2V}=1.5)$0.4571	756.6031	0.0166
600	3.5	30	$(c_{2V}=2.0)$1.9990	698.1602	0.0757
650	2.5	30	$(c_{2V}=4.0)$19.7177	663.8764	0.7653

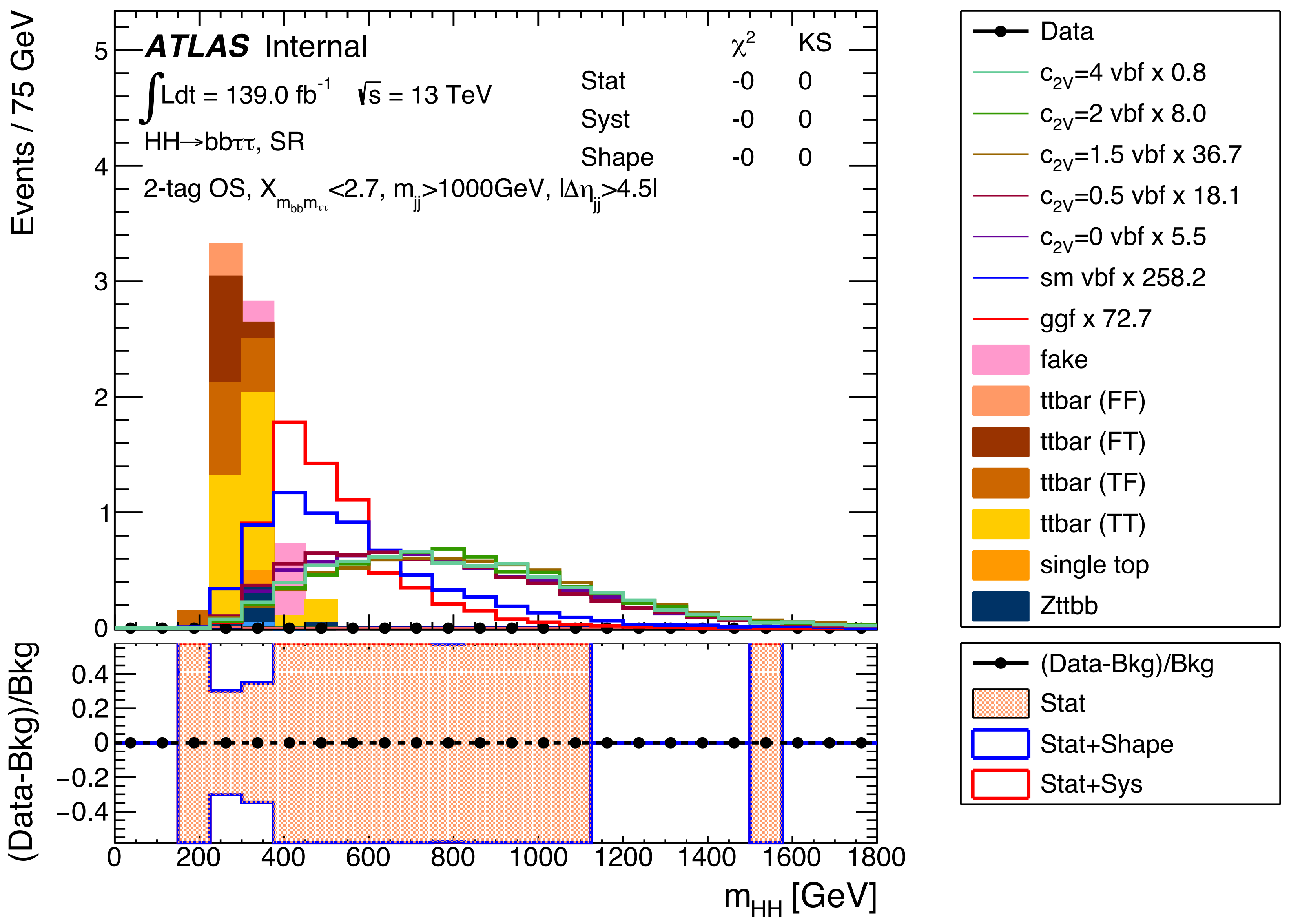
(only main backgrounds: Ztt, ttbar and fakes)

M_{jj} [GeV]	$\Delta\eta_{jj}$	$p_T^{j_1}$ [GeV]	signal yield	bkg yield	$(s/\sqrt{b})$
600	3.5	35	$(c_{2V}=0.0)$2.8610	636.4448	0.1134
600	3.5	35	$(c_{2V}=0.5)$0.8678	636.4448	0.0344
600	2.5	30	$(c_{2V}=1.5)$0.4557	710.9719	0.0171
600	3.5	30	$(c_{2V}=2.0)$1.9943	654.3321	0.0780
650	2.5	30	$(c_{2V}=4.0)$19.6611	623.7541	0.7872

With tau vertex requirement

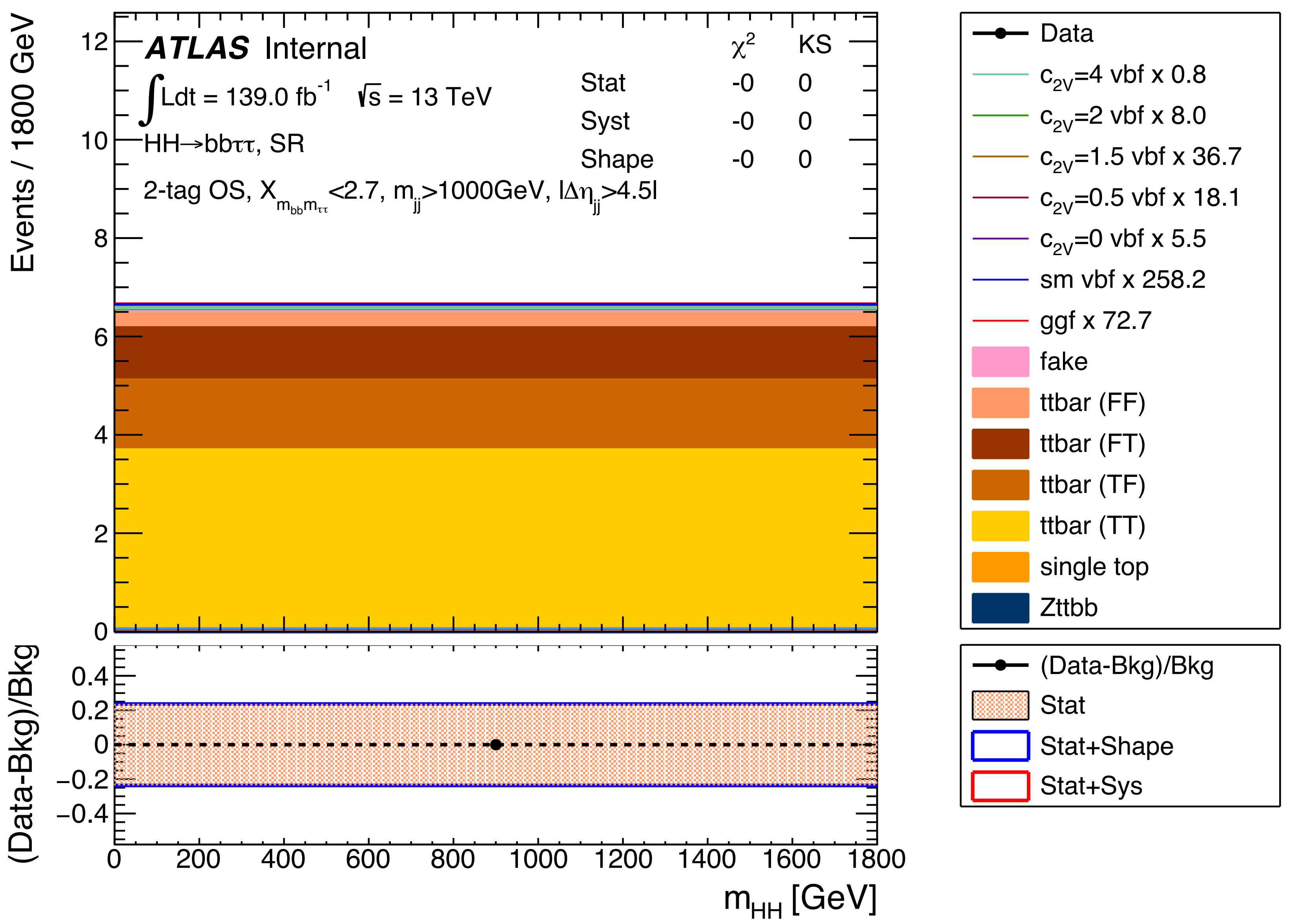
$\tau_{had}\tau_{had}$: Cut Based Studies for c_{2V} Scan

Low statistics



$\Rightarrow$

One bin fit



- $bb\tau\tau$ event selection
- VBF jet selection
- Cut on $X_{m_{bb}m_{\tau\tau}} < 2.7$ $\quad X_{m_{bb}m_{\tau\tau}} = \sqrt{\left(\frac{m_{bb} - 112\ \text{GeV}}{0.1 \cdot m_{bb}}\right)^2 + \left(\frac{m_{\tau\tau} - 116\ \text{GeV}}{0.1 \cdot m_{\tau\tau}}\right)^2}$
- Cut on $m_{jj} > 1000$ GeV and $|\Delta\eta_{jj}| > 4.5$

c_{2V}	-2σ	-1σ	μ_{VBF}	$+1\sigma$	$+2\sigma$
0.0	3.2	4.3	5.9	8.7	12.8
0.5	11.3	15.2	21.1	31.8	48.1
1.0	165.4	222.1	308.2	467.5	716.5
1.5	20.9	28.2	39.2	57.0	82.5
2.0	4.8	6.4	8.9	13.2	19.4
4.0	0.5	0.6	0.9	1.3	1.9

VHH Contamination

Corrected cross section calculation considering
VHH Contamination fractions f_{VBF} in each signal sample

$$\sigma_{\mathrm{VBF}}^{\mathrm{N3LO}}(c_{2V}{=}\mathrm{x}) = \sigma_{\mathrm{VBF}}^{\mathrm{N3LO}}(c_{2V}{=}1) \cdot \frac{\sigma^{\mathrm{LO}}(c_{2V}{=}\mathrm{x})}{\sigma^{\mathrm{LO}}(c_{2V}{=}1)} \cdot \frac{f_{\mathrm{VBF}}(c_{2V}{=}\mathrm{x})}{f_{\mathrm{VBF}}(c_{2V}{=}1)}$$

with $f_{\mathrm{VBF}}(c_{2V}{=}1) = 79.5\%$

c_{2V}	$f_{\mathrm{VBF}}(c_{2V}{=}\mathrm{x})$	$\sigma_{\mathrm{VBF}}^{\mathrm{N3LO}}(c_{2V}{=}\mathrm{x})$
0	99.55%	27.629 fb
0.5	97.99%	9.826 fb
1.5	86.58%	3.637 fb
2	94.49%	15.221 fb
4	98.09%	159.396 fb