VBS VVH All-Hadronic

Switching to XVqq

January 30th, 2023

*P. Chang, L. Giannini, **J. Guiang**, Y. Xiang, E. Zenhom*

Cut-based

Yields scaled to lumi×σ, rounded for readability

Cut	QCD	TTbarJets	TotalBkg	Eff.*	VBSVVH (C_{2V} = 2)	Eff.*
Skim	136,950K	748K	137,698K	—	175	—
HLT + MET Filters	88,614K	574K	89,188K	35%	168	4%
At least 3 fat jets	686K	16K	701K	99%	35	79%
Object selection	290K	10K	300K	57%	20	42%
M_{jj} > 500 GeV	87K	3.1K	90K	70%	18	12%
$\|\Delta\eta_{jj}\|$ > 500 GeV	81K	2.9K	84K	6%	18	0%
H→b$\bar{b}$ fat jet PNet Xbb > 0.9	4K	948	**4.9K**	94%	**12**	35%
V→qq fat jet PNet X > 0.9	21	17	**38**	99%	**5**	61%
S_T > 1300 GeV	10	11	**21**	45%	**4**	5%
M_{SD}(H→b$\bar{b}$ fat jet) < 150 GeV	3	6	**8**	61%	**4**	2%
M_{SD}(V→qq fat jet) < 120 GeV	0	2	**2**	75%	**4**	11%

Yields scaled to lumi×σ, rounded for readability

Cut	QCD	TTbarJets	TotalBkg	Eff.*	VBSVVH (C_{2V} = 2)	Eff.*
Skim	136,950K	748K	137,698K	—	175	—
HLT + MET Filters	88,614K	574K	89,188K	35%	168	4%
At least 3 fat jets	686K	16K	701K	99%	35	79%
Object selection	290K	10K	300K	57%	20	42%
M_{jj} > 500 GeV	87K	3.1K	90K	70%	18	12%
$\|\Delta\eta_{jj}\|$ > 500 GeV	81K	2.9K	84K	6%	18	0%
H→b$\bar{b}$ fat jet PNet Xbb > 0.9	4K	947	**4.9K**	94%	**12**	35%
V→qq fat jet PNet X > 0.9	13	9	**22**	100%	**4**	68%
S_T > 1300 GeV	7	5	**12**	42%	**3**	5%
M_{SD}(H→b$\bar{b}$ fat jet) < 150 GeV	1	2	**4**	70%	**3**	2%
M_{SD}(V→qq fat jet) < 120 GeV	0	1	**1**	64%	**3**	6%

- Left: using ParticleNet **XVqq** = (Xbb + Xcc + Xqq)/(Xbb + Xcc + Xqq + QCD)
- Right: using ParticleNet **XWqq** = (Xcc + Xqq)/(Xcc + Xqq + QCD)
- Not optimized, so maybe BDT makes better use of the XVqq score

*eff = 1 - after/before

BDT

Yields scaled to lumi×σ, rounded for readability

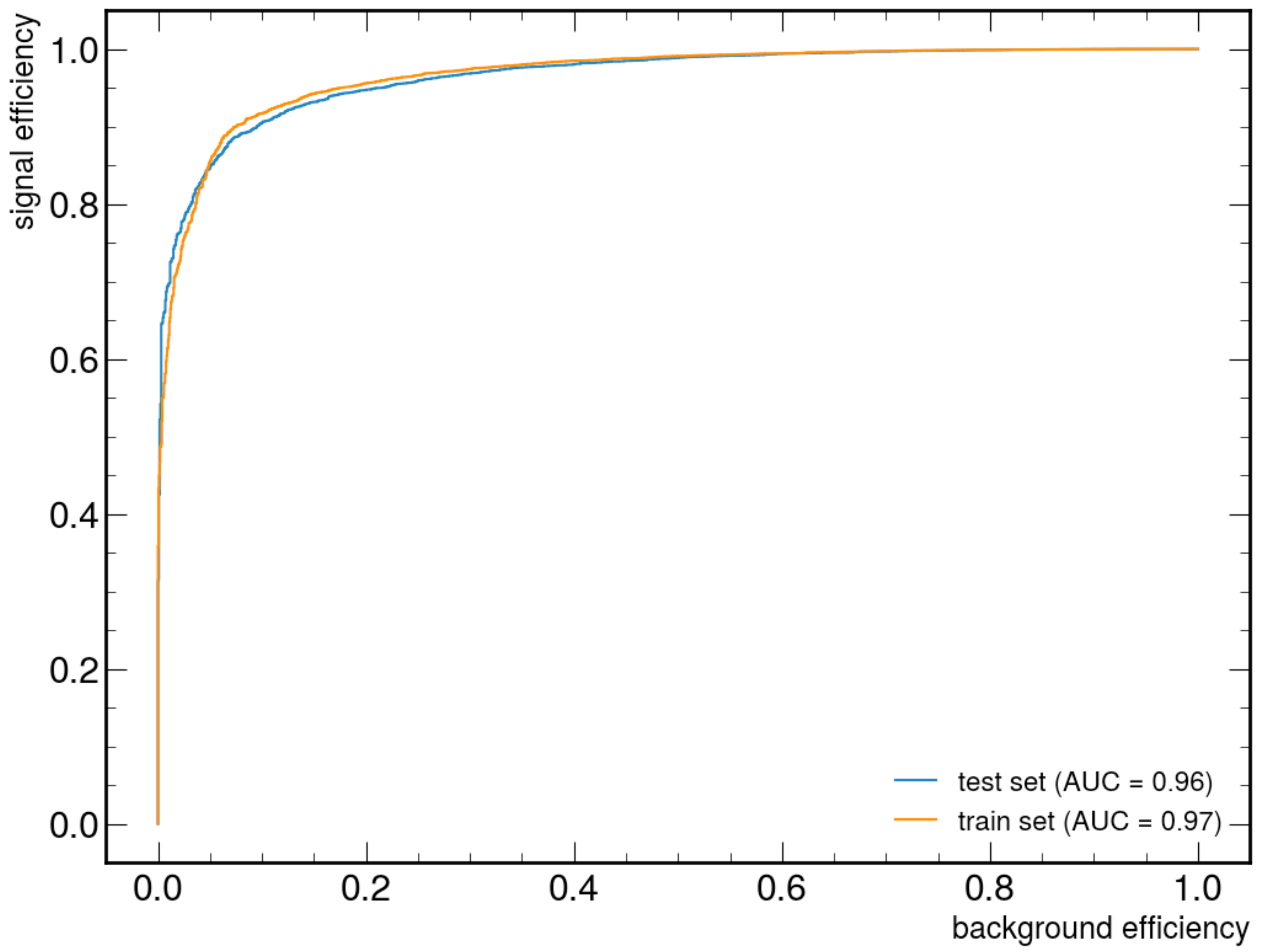
Cut	QCD	TTbarJets	TotalBkg	Eff.*	VBSVVH (C_{2V} = 3)	Eff.*
BDT Preselection	607	151	758	—	10	—

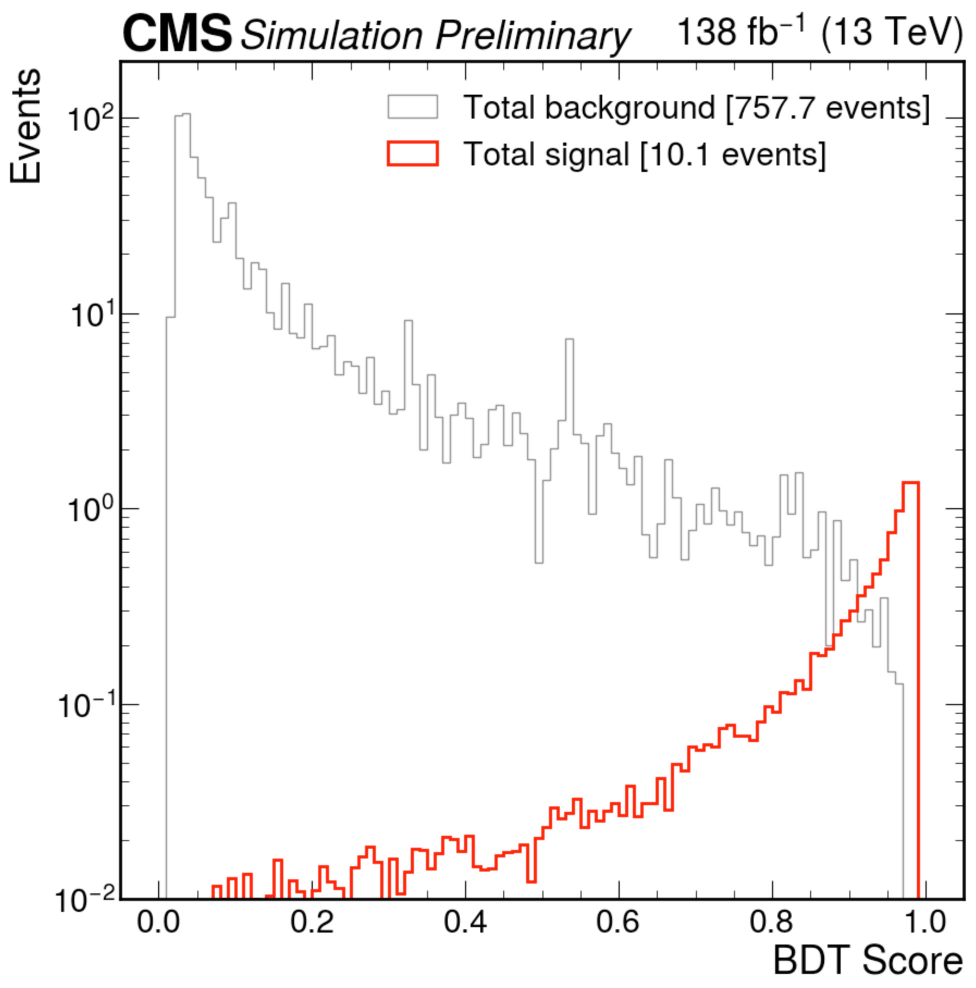
- First apply BDT preselection:
  - $S_T > 1300$ GeV
    AND $H\to b\bar{b}$ fat jet PNet Xbb > 0.5
    AND $V\to qq$ fat jet PNet **X** > 0.5
    AND $M_{SD}(H\to b\bar{b}$ fat jet$) < 150$ GeV
    AND $M_{SD}(V\to qq$ fat jet$) < 120$ GeV
- Train BDT using XGBoost with the input features listed on the right
  - 40/60 test/train split
  - Hyperparameters listed in backup

Object	Features
$H\to b\bar{b}$ fat jet	• ParticleNet Xbb score • p_T • M_{SD}
$V\to qq$ fat jets	• ParticleNet **X** score • p_T • M_{SD}
Other	• MET • N_{jets} (AK4)

BDT

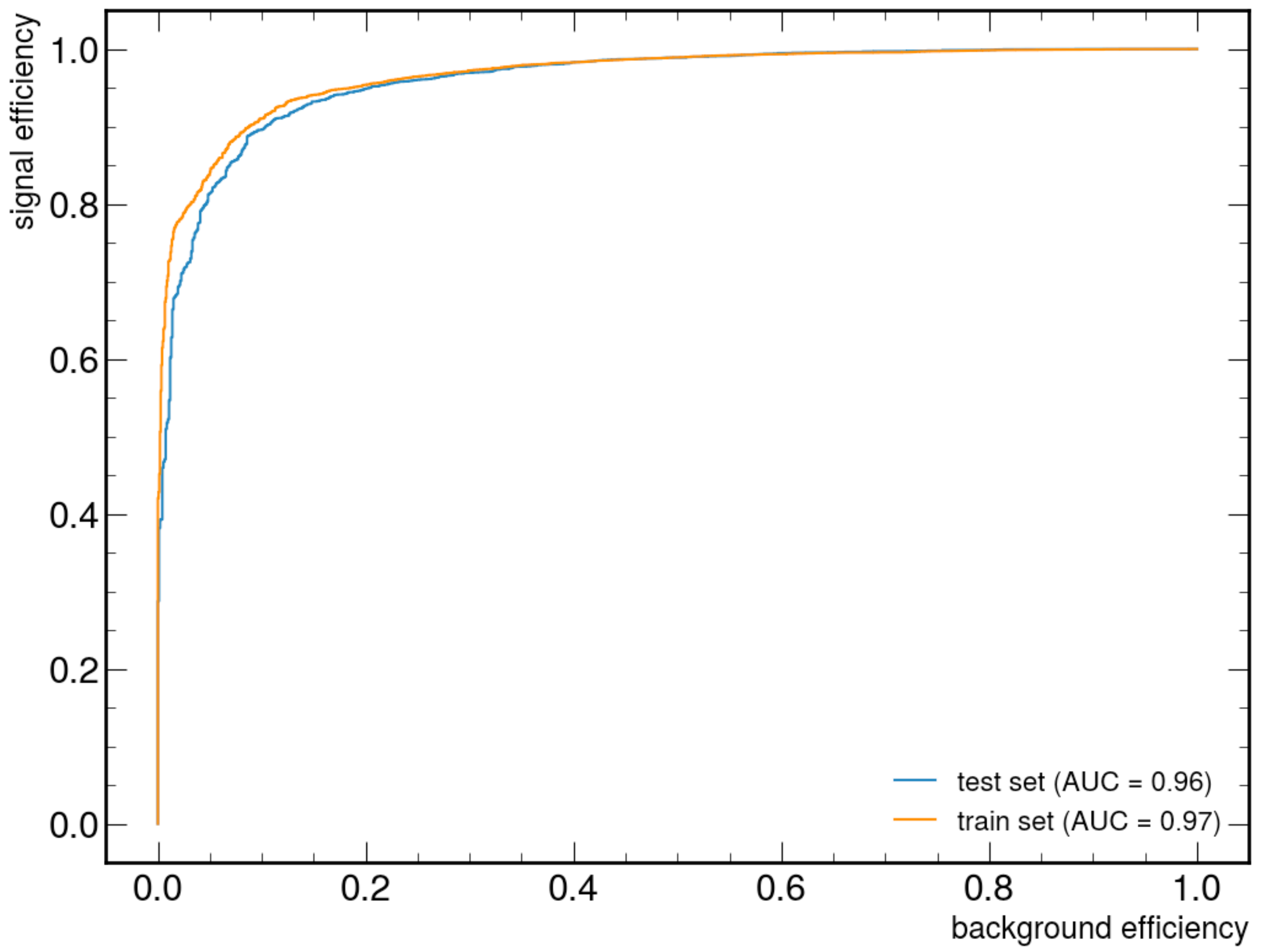
Using XVqq

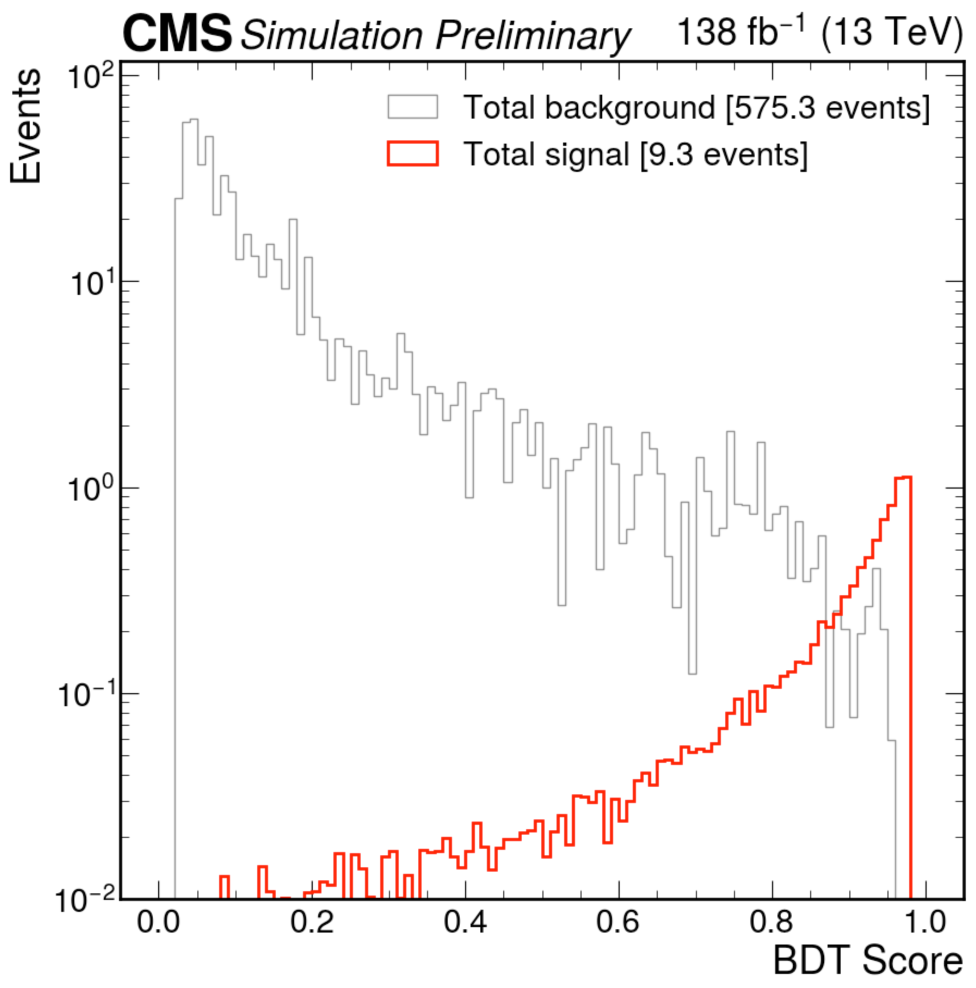




Selection	Signal (C_{2V} = 2)		Total Background		S/√B
	wgt.	raw	wgt.	raw	
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$	9.24	17063	253.16	947	0.58
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.80	7.46	13779	3.14	35	4.21
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.85	6.94	12800	1.43	20	5.80
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.86	6.78	12518	1.36	19	5.80
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.87	6.62	12227	1.08	15	6.36
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.88	6.45	11906	1.03	14	6.36
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.89	6.24	11524	0.56	9	8.35
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.90	6.00	11081	0.49	8	8.57

Using XWqq





Selection	Signal (C_{2V} = 2)		Total Background		S/√B
	wgt.	raw	wgt.	raw	
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$	8.51	15488	178.97	718	0.64
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.80	6.68	12131	1.75	20	5.05
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.85	6.11	11046	0.93	12	6.35
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.86	5.95	10770	0.79	10	6.70
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.87	5.75	10402	0.64	8	7.17
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.88	5.56	10052	0.57	7	7.33
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.89	5.33	9645	0.53	6	7.33
$M_{jj} > 500$ and $\|\Delta\eta_{jj}\| > 3$ and BDT > 0.90	5.06	9165	0.46	5	7.45

Difference is not huge, but slightly better using XVqq

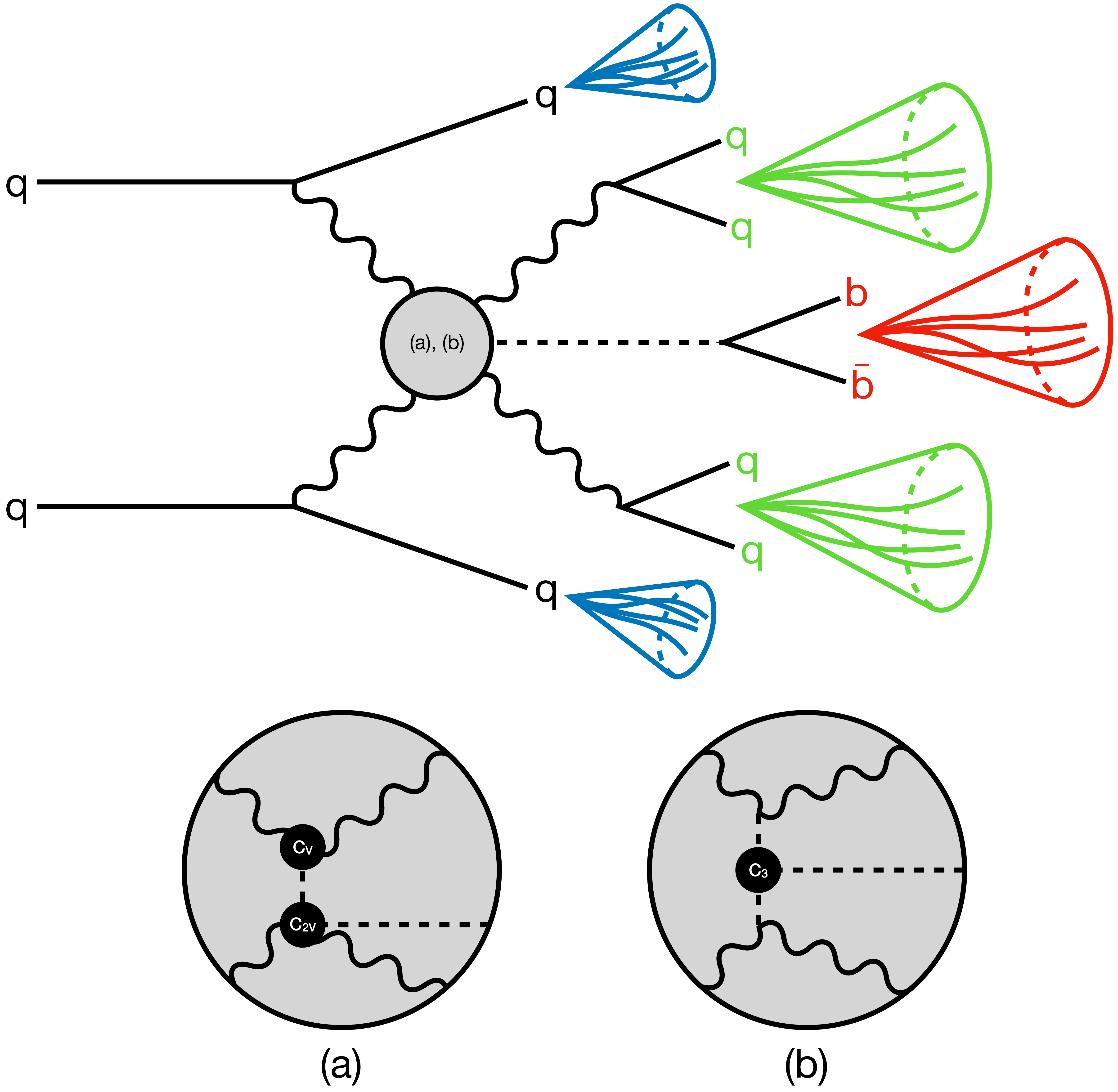
Summary

- Now select ≥ 3 fat jets (rather than == 3 fat jets)
- Compared XVqq vs. XWqq
  - Cut-based is worse with XVqq (but not optimized)
  - BDT-based is better with XVqq

Backup

All-Hadronic VBS VVH

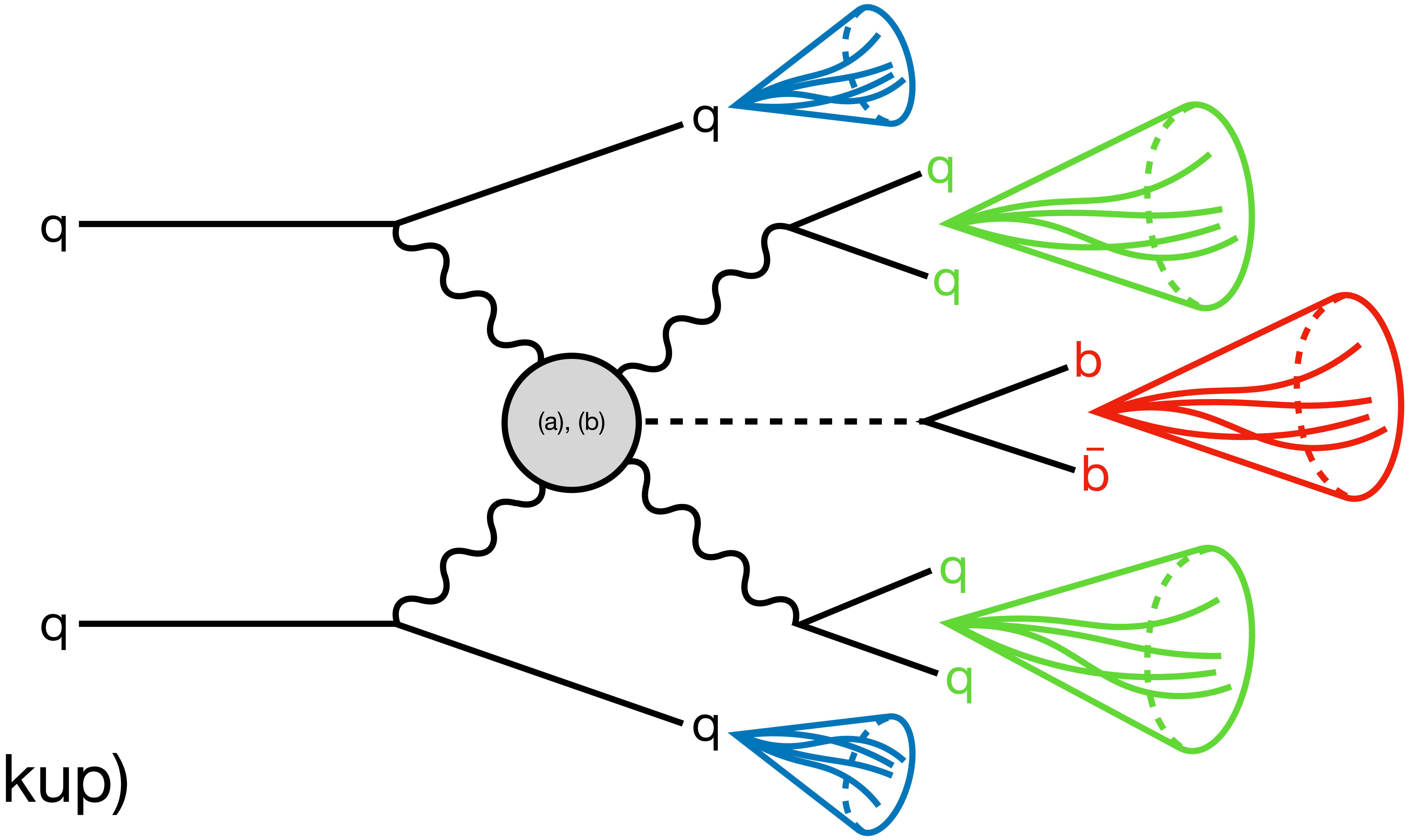
- Targeting the following final states:
  - WWH → qq qq qq $b\bar{b}$
  - ZZH → qq qq qq $b\bar{b}$
  - WZH → qq qq qq $b\bar{b}$
- Sensitive to C_{2V}, C_3, and C_V in principle
- BSM signature:
  - W/Z/H jets with large p_T
  - VBS jets with large $\Delta\eta_{jj}$, M_{jj}

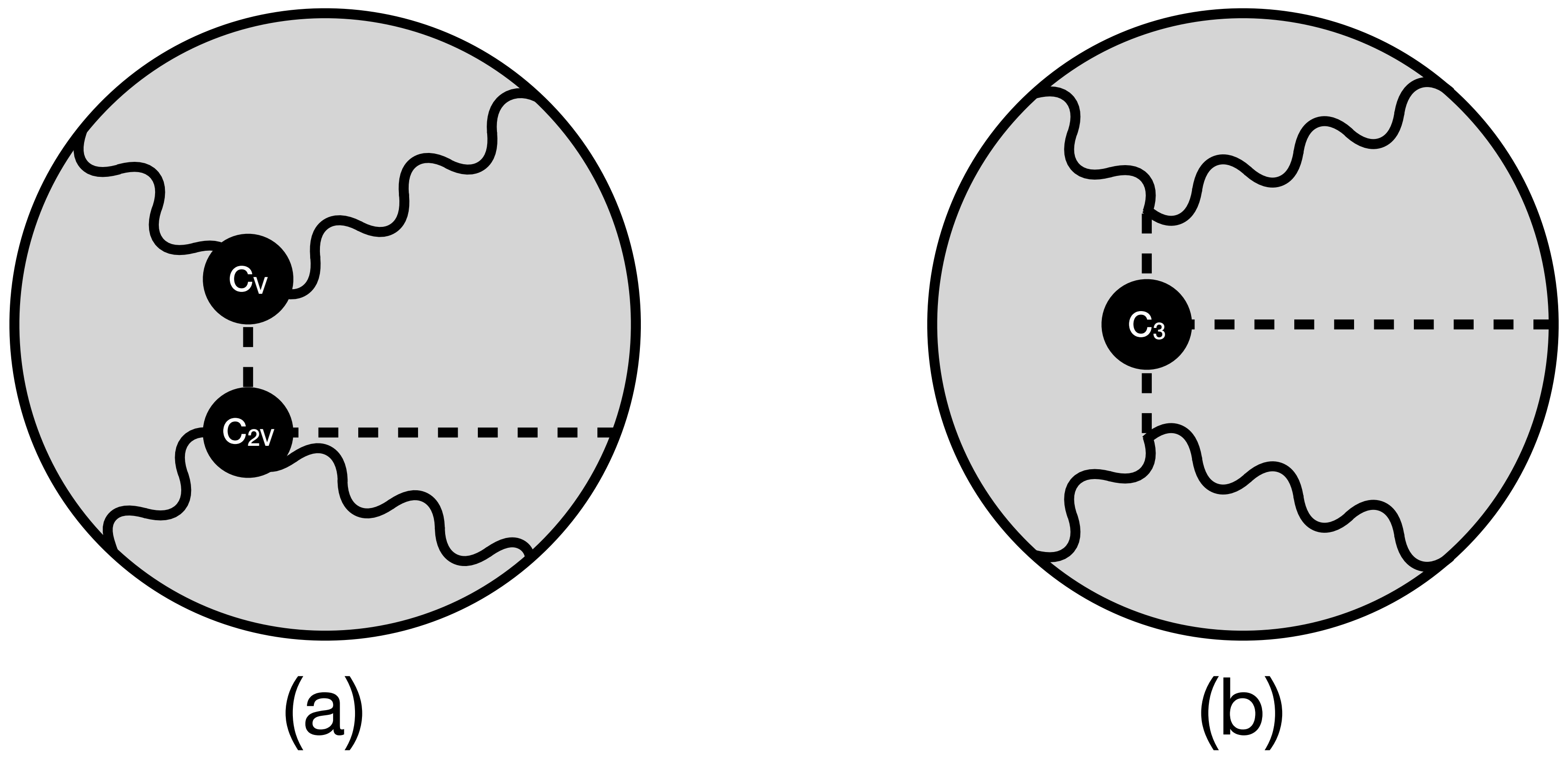


UC San Diego

All-Hadronic VBS VVH

- One interesting N_{jets} vs. $N_{fatjets}$ channel:
    - **3 AK8 fat jets, ≥ 2 AK4 jets** (right)
    - ~~2 AK8 fat jets, ≥ 4 AK4 jets~~
    - ~~2 AK8 fat jets, 3 AK4 jets~~
- Practically zero signal with $N_{fatjets}$ > 3 (backup)
- From previous studies, $N_{fatjets}$ == 1 channel not worthwhile pursuing right now





Analysis Skim

Yields scaled to lumi×σ, rounded for readability

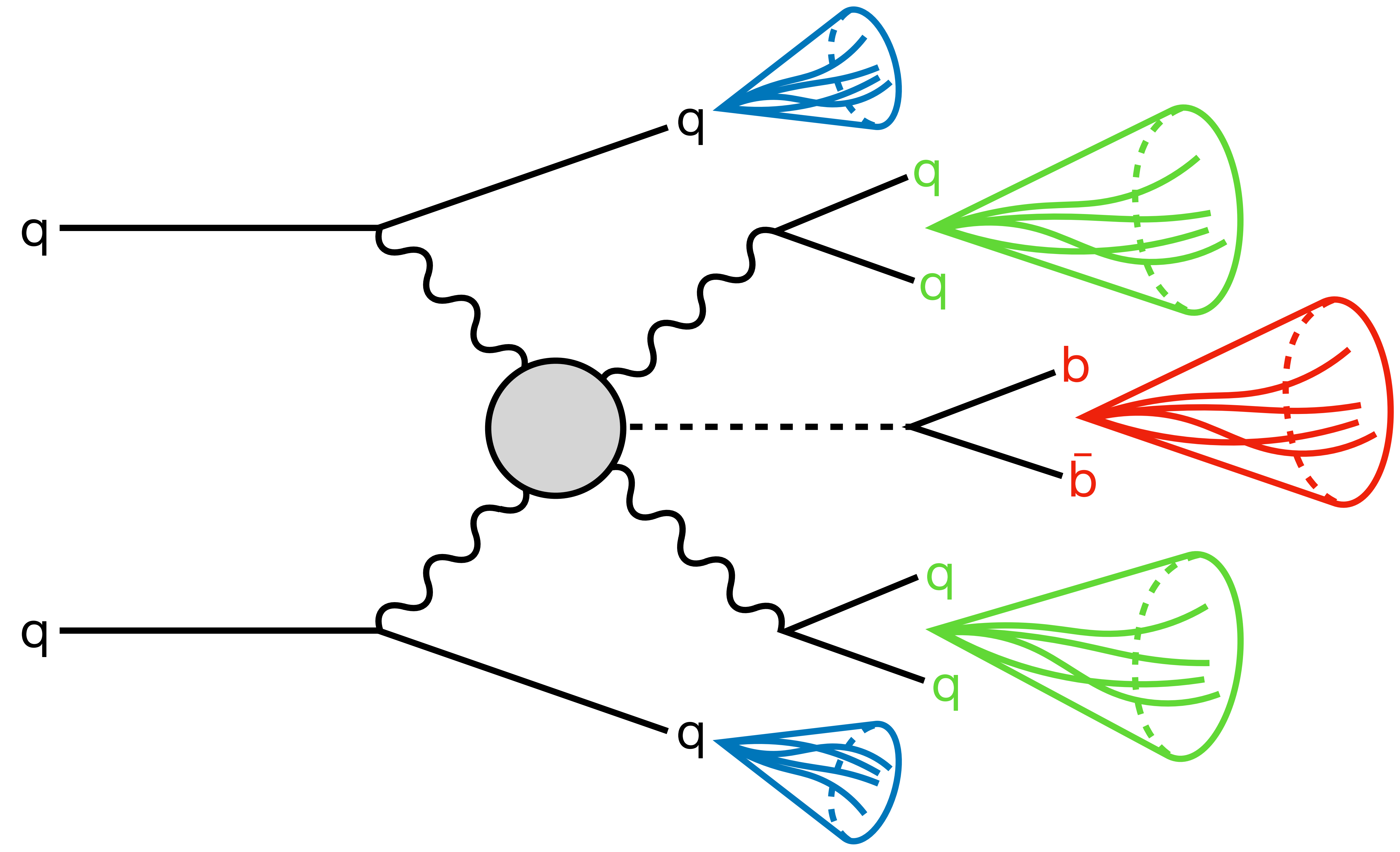
Cut	QCD	TTbarJets	TotalBkg	Eff.*	VBSVVH (C_{2V} = 2)	Eff.*
Skim	136,950K	748K	137,698K	—	175	—

Object	Selection
Leptons (μ, e)	== 0 veto*
Fat Jets	≥ 2 w/ p_T > 300 GeV AND $\|\eta\|$ < 2.5 AND mass > 50 GeV AND M_{SD} > 40 GeV AND fat jet ID > 0
Jets	≥ 2 w/ p_T > 20 GeV AND passes tight jet ID AND ΔR(jet, fat jet) > 0.8
Other	≥ 1 pair of AK4 jets w/ M_{jj} > 500 GeV AND $\|\Delta\eta_{jj}\|$ > 3

*Using the ttH lepton ID

UC San Diego

Object Selections

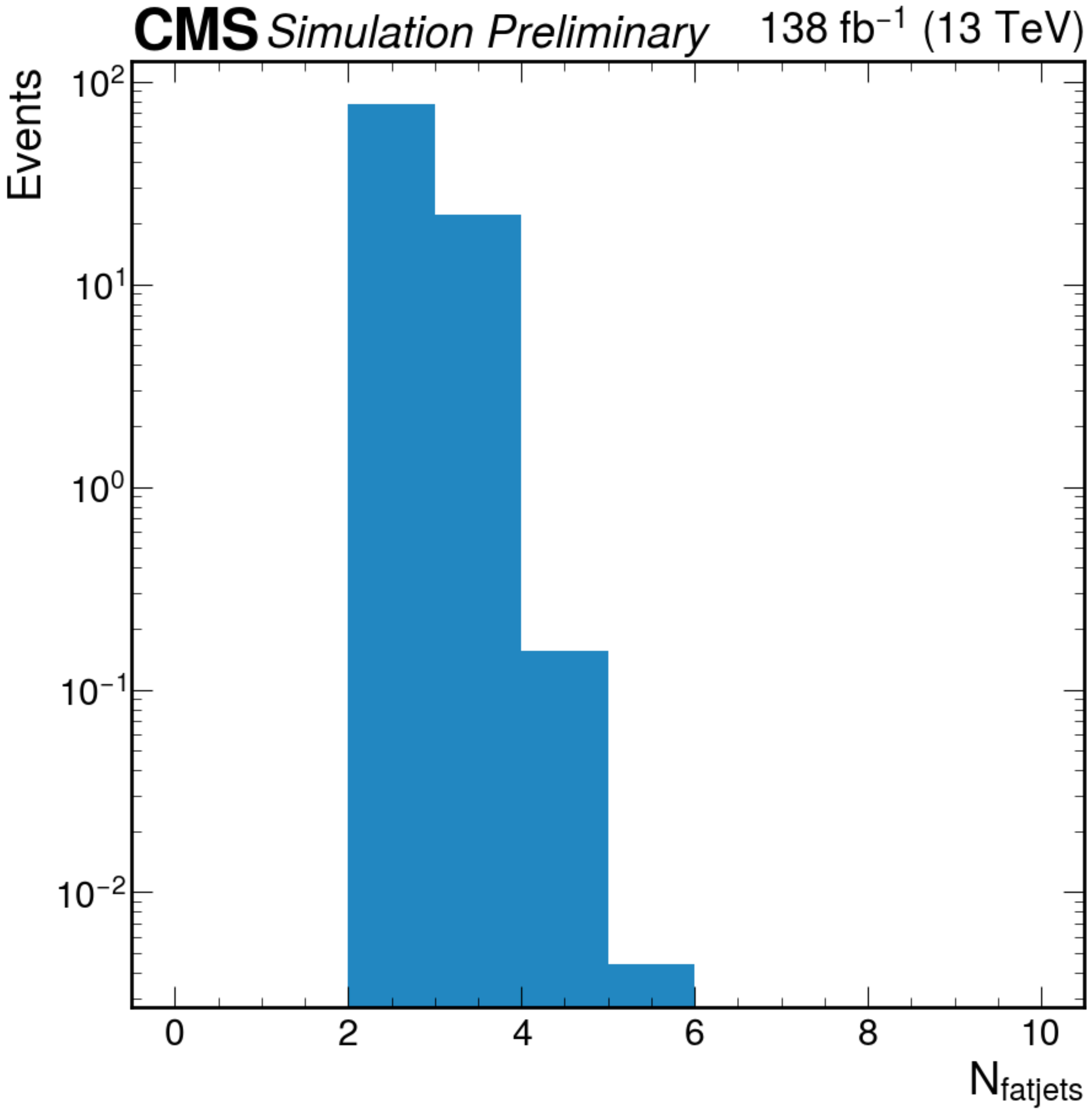


Object	Selections
AK8 jets (same as skim)	• p_T > 300 GeV • Fat jet ID > 0 • $\|\eta\|$ < 2.5 • Mass > 50 GeV • M_{SD} > 40 GeV
H→b$\bar{b}$ fat jet	• Has max(ParticleNet Xbb)
V→qq fat jets	• Leading in ParticleNet XVqq
AK4 jets (same as skim)	• p_T > 20 GeV • Passes tight jet ID • ΔR(AK4, AK8) > 0.8
VBS (AK4) jets	• p_T > 30 GeV • For > 2 candidates**: • Take pair with maximum $\|\Delta\eta_{jj}\|$

*ParticleNet XWqq = mass-decorrelated, W-to-qq-like tagger
**Used in Yanxi's analysis, but can easily be changed to sync with Yifan

$N_{fatjets}$

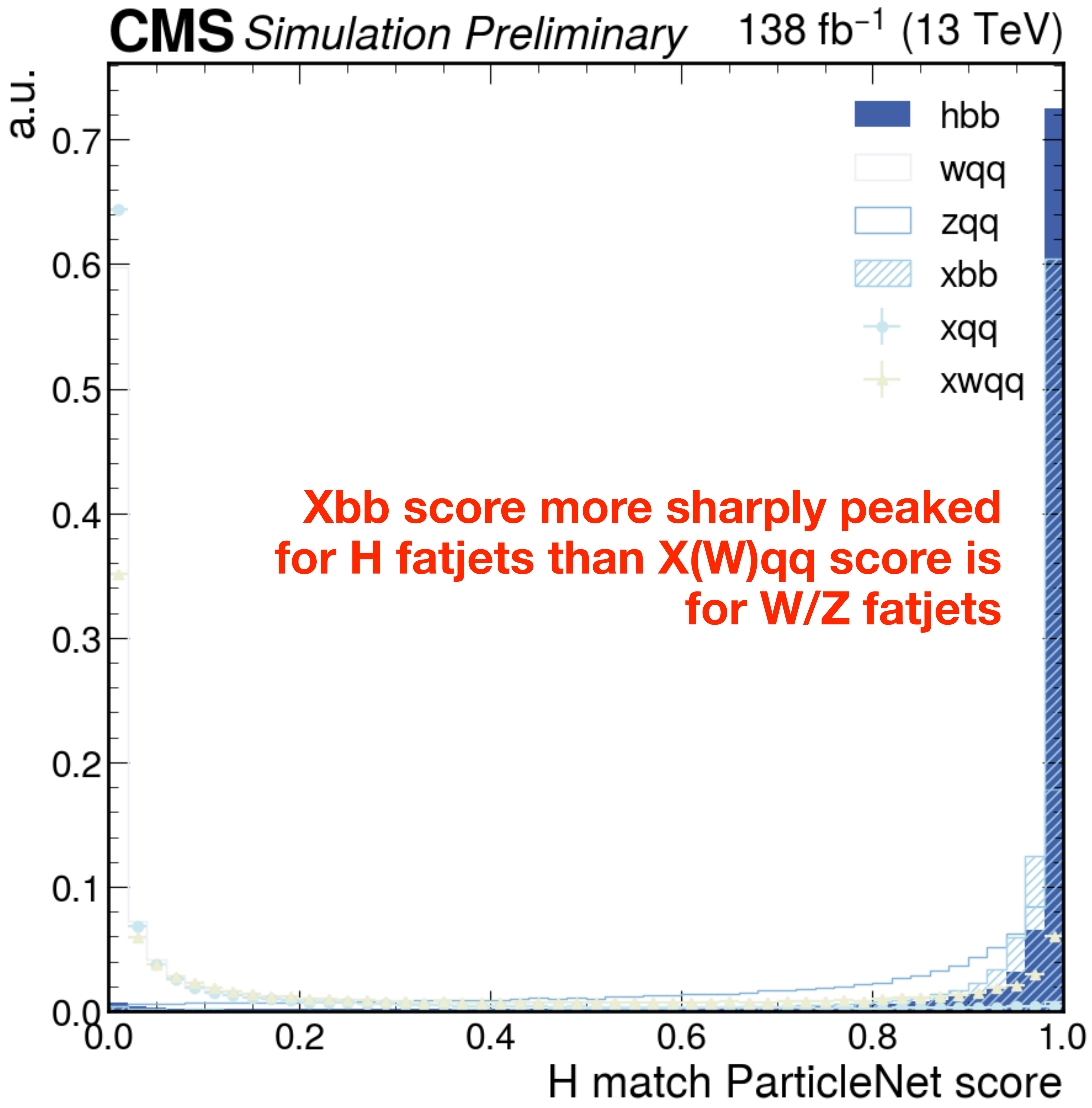
- Plotting sum of all VBS VVH signals here
- Selections applied: skim, HLT triggers, gen-level H, W, Z decay hadronically
- Practically 0 events with more than three fatjets



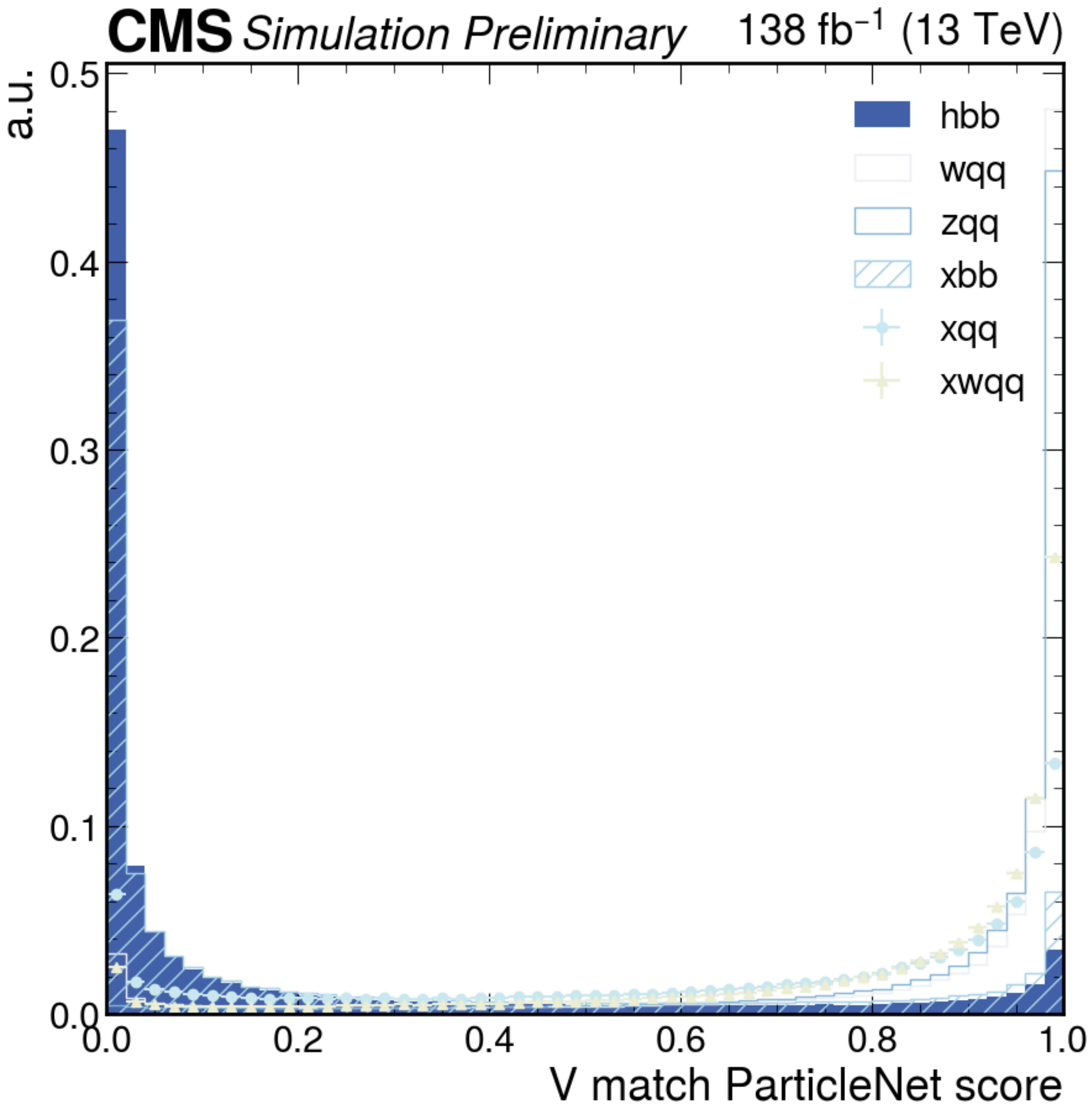
CMS

UC San Diego

ParticleNet Taggers

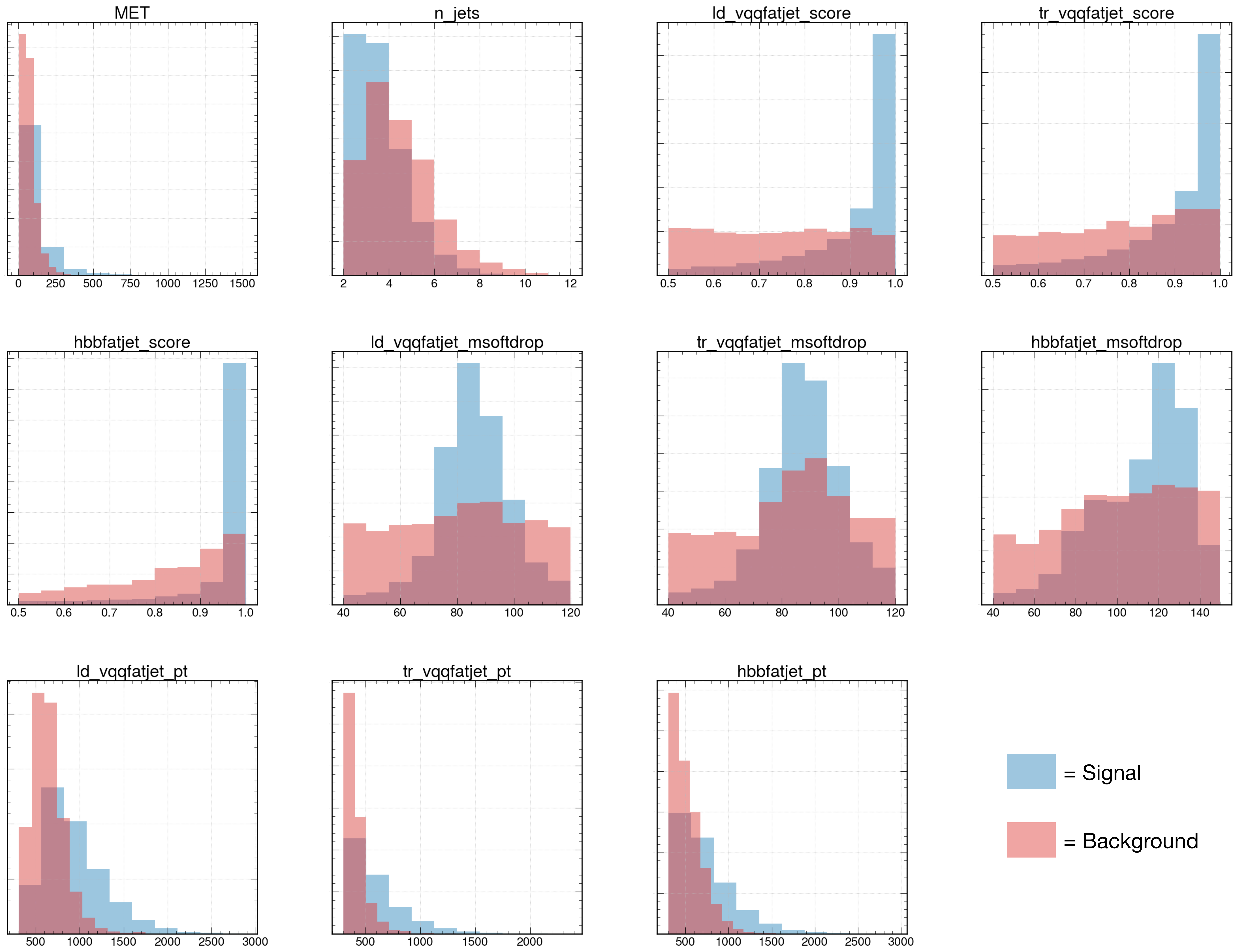


Xbb score more sharply peaked for H fatjets than X(W)qq score is for W/Z fatjets



Mass-correlated taggers more sharply peaked, but MD taggers may be sufficient

BDT



Parameter	Value	Description*
`objective`	binary:logistic	Learning objective; ‘binary:logistic’ specifies logistic regression for binary classification, output probability
`eta`	0.1	Step size shrinkage (alias: learning_rate)
`max_depth`	3	Max. depth of tree: larger = more complex = more prone to overfitting
`verbosity`	1	0 (silent), 1 (warning), 2 (info), 3 (debug)
`nthread`	8	Number of parallel threads
`eval_metric`	auc	Evaluation metrics for validation data. 'auc' = Area Under the Curve
`subsample`	0.6	Subsample ratio of the training instances
`alpha`	8.0	L1 regularization term on weights: Larger = more conservative
`gamma`	2.0	Min. loss rediction to make leaf (alias: min_split_loss)
`lambda`	1.0	L2 regularization term on weights: Larger = more conservative
`min_child_weight`	1.0	Minimum sum of instance weight (hessian) needed in a child
`colsample_bytree`	1.0	The subsample ratio of columns when constructing each tree
`scale_pos_weight`	75.0	Control the balance of positive and negative weights, useful for unbalanced classes

*From: https://xgboost.readthedocs.io/en/stable/parameter.html

H_T-binned QCD Samples

- Ignoring the $H_T \in [50, 100)$ GeV sample due to low stats and irrelevant anyway
- Cross sections taken from AN-21-045
  - Measurement of the Higgs boson production via Vector Boson Fusion process with subsequent decay of the Higgs boson into a pair of bottom quarks

Table 3: List of simulated samples along with the rates used in this analysis, where [*] denotes the ultra legacy production campaign of the simulated samples, for 2016-APV, 2016-NonAPV and 2018.

Process	Sample	$\sigma \times BR$ [pb]
VBFH $\rightarrow b\bar{b}$	VBFHToBB_M-125_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	2.233
	VBFHToBB_M-125_dipoleRecoilOn_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	
	/VBFHToBB_M-125_TuneCH3_13TeV-powheg-herwig/[*]/MINIAODSIM	
ggH$\rightarrow$bb	/GluGluHToBB_M-125_TuneCP5_13TeV-amcatnloFXFX-pythia8/[*]/MINIAODSIM	28.293
	/GluGluHToBB_M-125_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	
QCD multijet in H_T bins	QCD_HT100to200_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	27849880
	QCD_HT200to300_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	1716997
	QCD_HT300to500_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	351302
	QCD_HT500to700_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	31630
	QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	6802
	QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	1206
	QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	98.71
	QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	20.2
$t\bar{t} + X$	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	88.29
	TTToHadronic_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	377.96
	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	365.34
single top	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	34.91
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/[*]/MINIAODSIM	34.40
	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/[*]/MINIAODSIM	115.30
	ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/[*]/MINIAODSIM	69.09
QCD Z + jets	ZJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	973.70
	ZJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	110.78
	ZJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	25.34
	ZJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	19.00
EWK Z + jets	EWKZ2Jets_ZToQQ_TuneCP5_13TeV-madgraph-pythia8/[*]/MINIAODSIM	9.92
QCD W + jets	WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	2549.00
	WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	276.50
	WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	59.25
	WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/[*]/MINIAODSIM	28.75

VBSVVH All-Hadronic Cutflow

	QCD		TTbarJets		TotalBkg		VBSVVH	
cut	raw	wgt	raw	wgt	raw	wgt	raw	wgt
Bookkeeping	23348898	136950279.23	11563843	747778.20	34912741	137698057.43	323711	175.10
SaveSystWeights	23348898	136950279.23	11563843	747778.20	34912741	137698057.43	323711	175.10
PassesEventFilters	23228634	136757063.13	11539080	746223.20	34767714	137503286.33	320310	173.29
PassesTriggers	19971141	88614027.81	8948283	574450.50	28919424	89188478.31	314509	168.28
SelectLeptons	19971141	88614027.81	8948283	574450.50	28919424	89188478.31	314509	168.28
NoLeptons	19971141	88614027.81	8948283	574450.50	28919424	89188478.31	314509	168.28
SelectFatJets	19971141	88614027.81	8948283	574450.50	28919424	89188478.31	314509	168.28
Geq3FatJets	718294	685582.18	235834	15550.80	954128	701132.98	64241	34.92
AllMerged_SelectVVHFatJets	718294	685582.18	235834	15550.80	954128	701132.98	64241	34.92
AllMerged_SelectJets	718294	685582.18	235834	15550.80	954128	701132.98	64241	34.92
AllMerged_SelectVBSJets	355043	289829.53	157726	10373.88	512769	300203.41	37585	20.41
AllMerged_SaveVariables	355043	289829.53	157726	10373.88	512769	300203.41	37585	20.41
AllMerged_STGt800	355043	289829.53	157726	10373.88	512769	300203.41	37585	20.41
AllMerged_MjjGt500	116277	86740.41	47932	3107.13	164209	89847.55	33177	18.00
AllMerged_detajjGt3	103848	81415.71	44921	2910.10	148769	84325.81	33048	17.93
AllMerged_XbbGt0p9	5233	3984.96	14657	948.15	19890	4933.12	21957	11.62
AllMerged_XWqqGt0p9	23	20.97	269	17.32	292	38.29	8640	4.52
AllMerged_STGt1300	18	9.86	173	11.08	191	20.94	8175	4.29
AllMerged_HbbMSDLt150	9	2.68	88	5.51	97	8.19	8030	4.21
AllMerged_VqqMSDLt120	1	0.20	29	1.84	30	2.05	6982	3.75

VBSVVH All-Hadronic Triggers

Year	HLT paths
2016	HLT_PFHT800 HLT_PFHT900 HLT_AK8PFHT650_TrimR0p1PT0p03Mass50 HLT_AK8PFHT700_TrimR0p1PT0p03Mass50 HLT_AK8PFJet450 HLT_AK8PFJet360_TrimMass30 HLT_AK8DiPFJet280_200_TrimMass30 HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20
2017	HLT_PFHT1050 HLT_AK8PFHT800_TrimMass50 HLT_PFJet320 HLT_PFJet500 HLT_AK8PFJet320 HLT_AK8PFJet500 HLT_AK8PFJet400_TrimMass30 HLT_AK8PFJet420_TrimMass30
2018	HLT_PFHT1050 HLT_AK8PFHT800_TrimMass50 HLT_PFJet500 HLT_AK8PFJet500 HLT_AK8PFJet400_TrimMass30 HLT_AK8PFJet420_TrimMass30