VBS WH Analysis

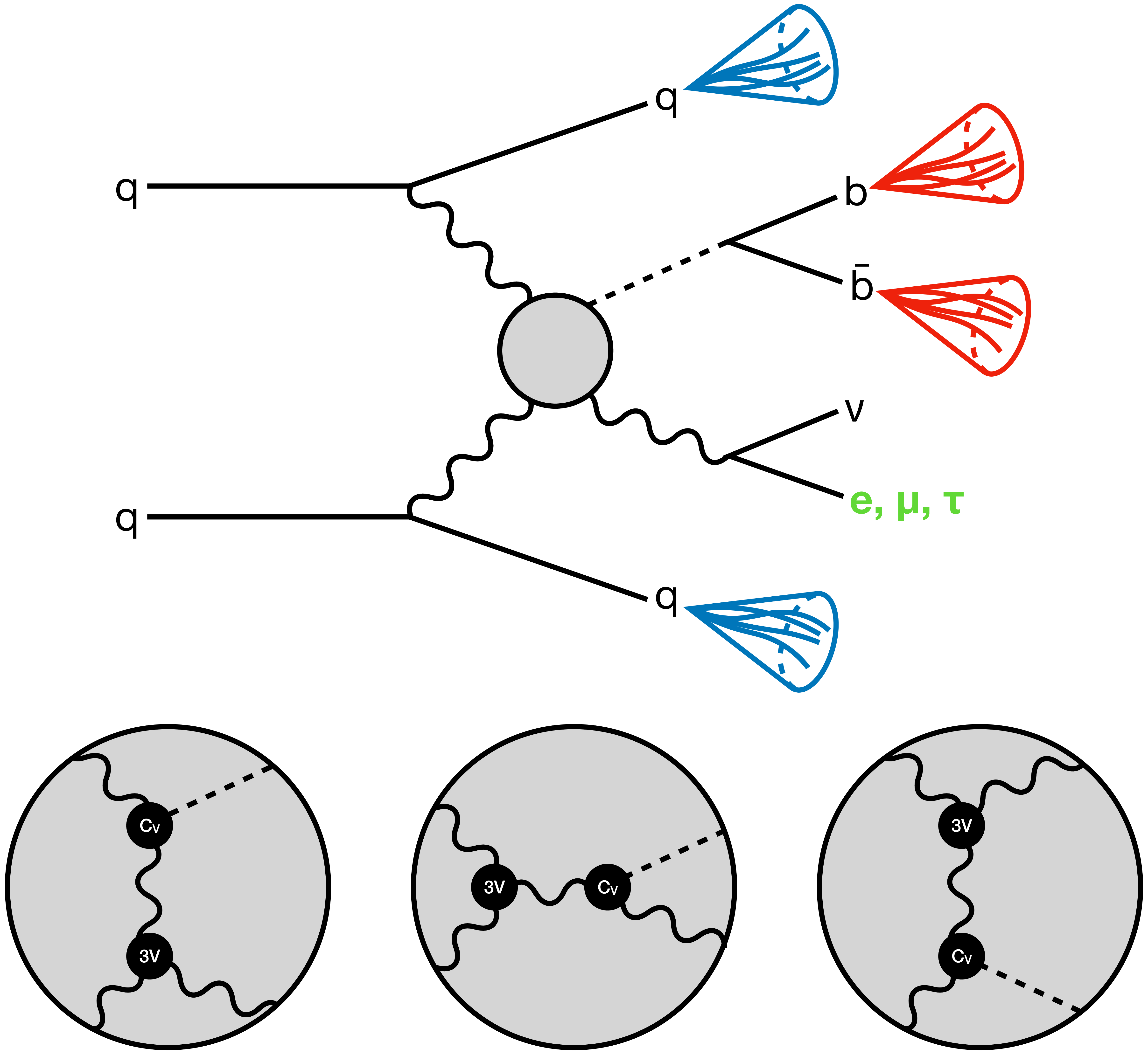
Towards excluding $\lambda_{WZ} < 0$

June 15th, 2022

P. Chang, ***J. Guiang***, *Y. Xiang*

Target Final State

- Targeting **VBS** WH→ℓν**b$\bar{b}$**
- Sensitive to κ_V → $\lambda_{WZ} = \kappa_Z/\kappa_W$
    - Handle for ruling out $\lambda_{WZ} < 0$ (BSM)
- VBS WH BSM kinematics:
    - High-p_T H and W (high S_T)
    - VBS jets with large $\Delta\eta_{jj}$, M_{jj}



(a) (b) (c)

CMS

UC San Diego

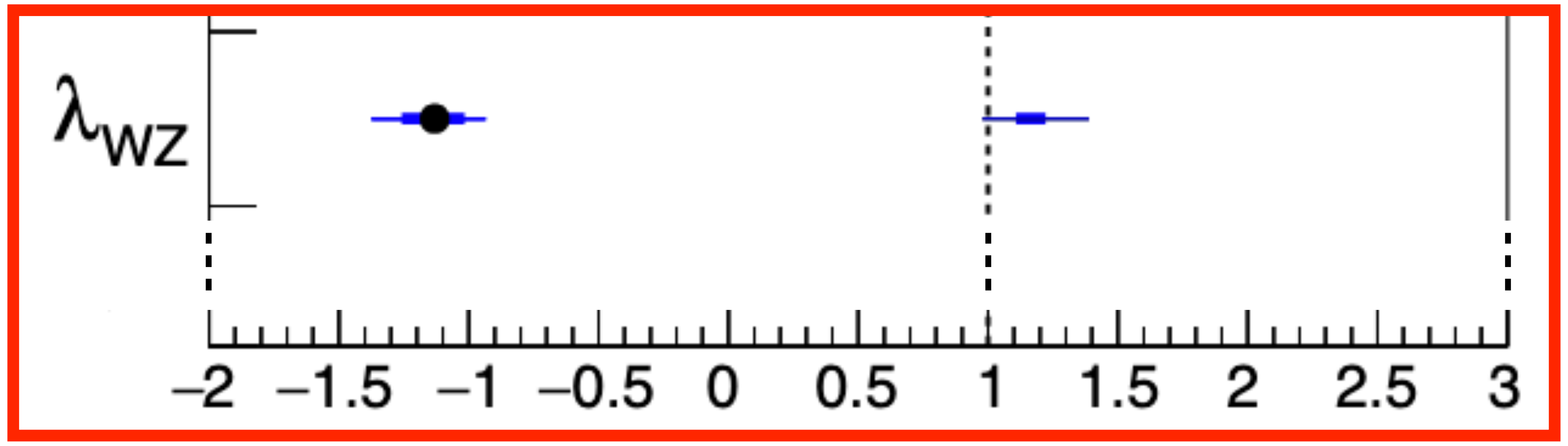
VBS WH Cross Sections

Model	σ [pb]
SM	0.075
$-\kappa_W$	0.433
$-\kappa_Z$	0.433

- Setting $\kappa_W < 0$ or $\kappa_Z < 0$ equivalently enhances cross section by an order of magnitude
- These numbers are taken from MadGraph*: `generate p p > w h j j QCD=0`
  - Includes gen-level filters (e.g. jet $p_T > 10$ GeV)
  - Generated 10,000 events for each to obtain xsec value
- Optimizing for $-\kappa_W$ (kinematics are equivalent to $-\kappa_Z$)
  - Generated 100k UL NanoAOD events for 2016 pre-VFP, 2016 post-VFP, 2017, and 2018

*More details in backup

Current CMS Result



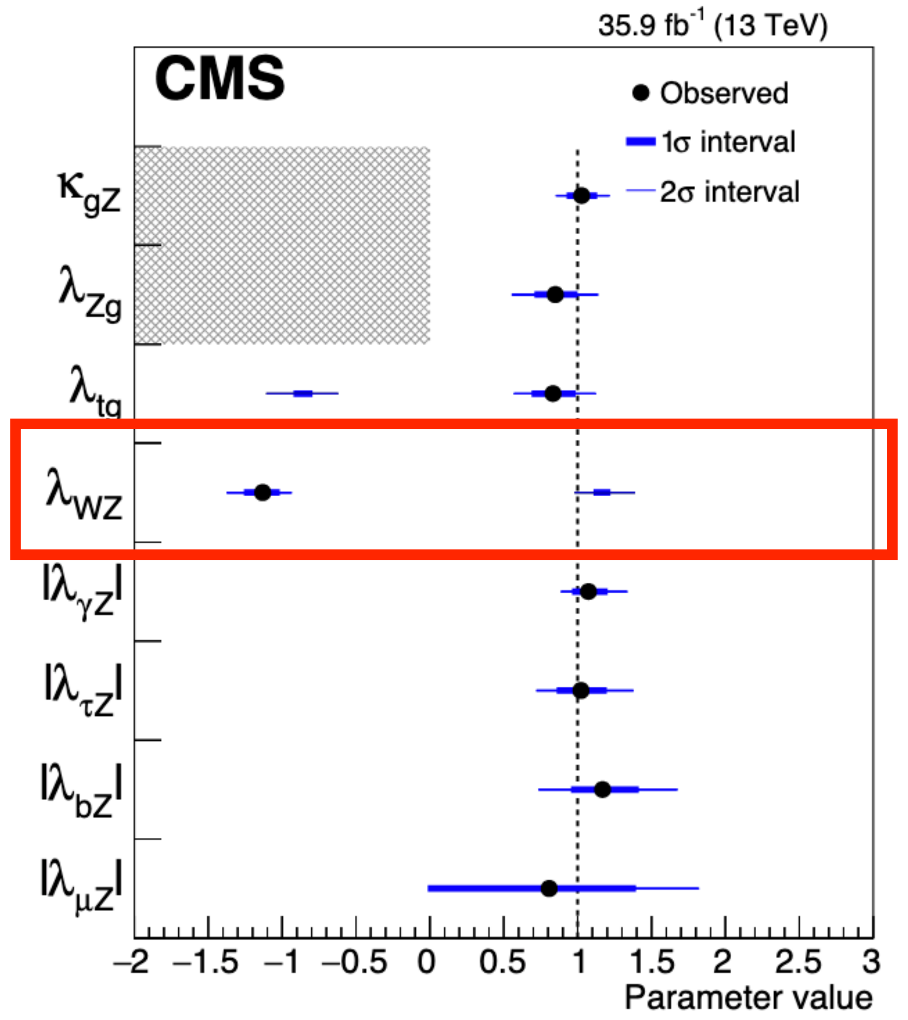


Figure 16: Summary of the model with coupling ratios and effective couplings for the ggH and $H \to \gamma\gamma$ loops. The points indicate the best fit values while the thick and thin horizontal bars show the 1σ and 2σ CL intervals, respectively. For this model, both positive and negative values of λ_{WZ} and λ_{tg} are considered.

- From 35.9 fb^{-1} Higgs combination: best fit hints at $\lambda_{WZ} = -1$ (BSM)
  - https://arxiv.org/pdf/1809.10733.pdf
- **With the handles we have on BSM λ_{WZ} via VBS WH, we can rigorously explore this!**

VBS WH Skims

Object	Primary Skim	Post-skim
Leptons (μ, e)	≥ 1 veto	== 1 loose AND == 1 tight
Fat Jets	≥ 1 w/ p_T > 200 GeV AND mass > 10 GeV AND M_{SD} > 10 GeV AND ΔR(fat jet, any veto lepton) > 0.8	≥ 1 w/ p_T > 250 GeV AND mass > 25 GeV AND M_{SD} > 25 GeV AND Hbb vs. QCD > 0.5 AND ΔR(fat jet, tight lepton) > 0.8
Jets	≥ 1 w/ p_T > 20 GeV AND ΔR(jet, any veto lepton) > 0.4	—

UC San Diego

VBS WH Preselection

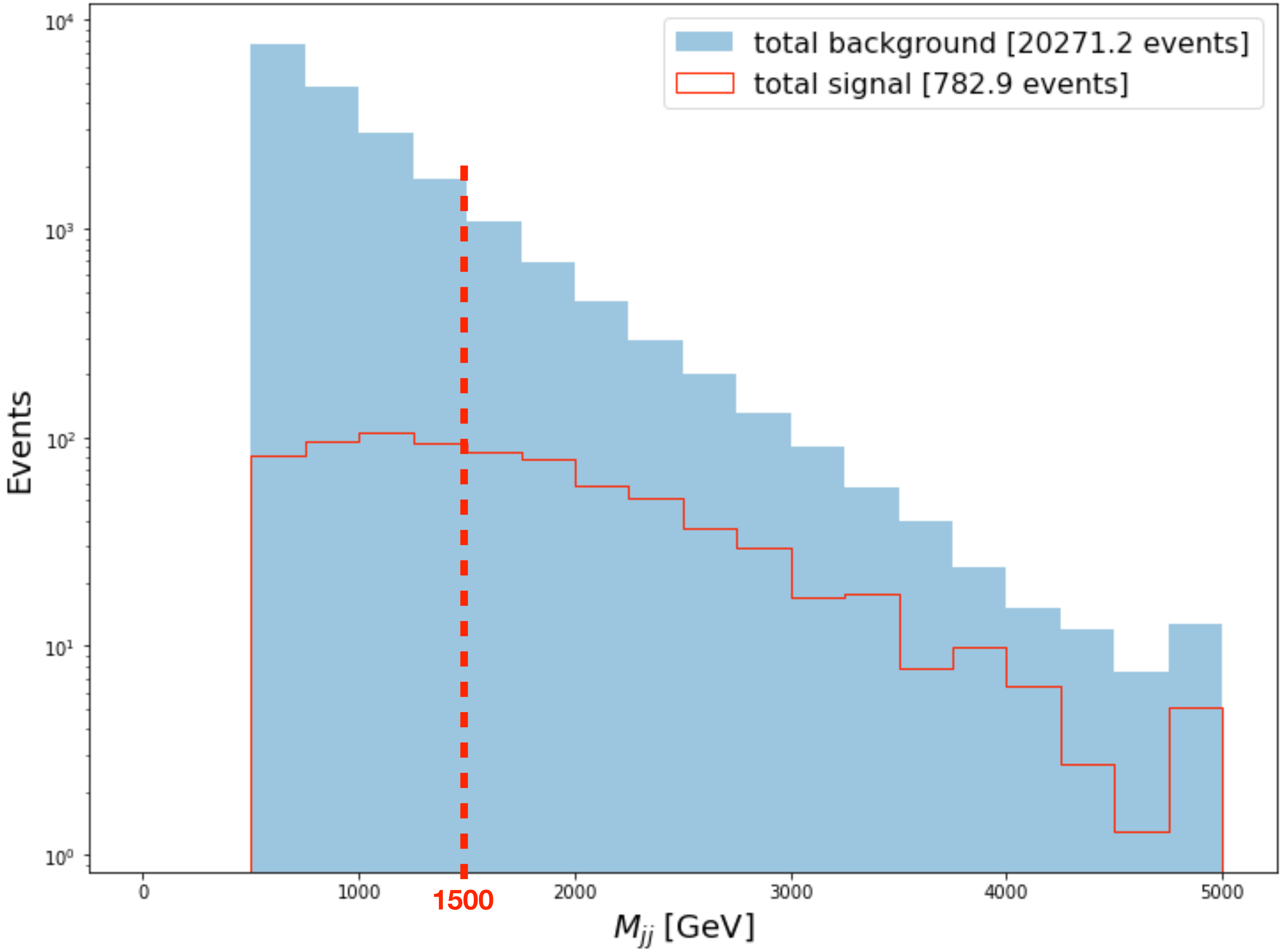
Cut	VHToNonbb	VV/VVV/VBSWZ	W+Jets	TTbar+X	TTbar1L	TTbar2L	TotalBkg	Eff.*	VBSWH (kW < 0)	Eff.*
Skim AND Postskim	167.20	1682.81	32660.21	3675.78	424281.50	34177.24	496644.74	—	59598.09	—
$p_T(\ell)$ > 40 GeV	134.51	1328.09	26056.05	2744.47	328311.10	25411.25	383985.47	23%	9071.35	85%
≥ 1 fat jet* w/ p_T > 250 GeV AND mass > 50 GeV AND M_{SD} > 40 GeV	134.75	1310.98	25501.02	2707.30	326174.50	23568.50	379397.05	1%	3281.05	64%
max(PNet Hbb vs. QCD) > 0.9	52.86	692.78	5851.27	1005.49	114319.10	9549.23	131470.73	65%	1344.32	59%
2 VBS jets (max E)**	16.38	220.91	3224.51	898.96	86676.40	6690.97	97728.13	26%	929.32	31%
M_{jj} > 500 AND $\Delta\eta_{jj}$ > 3	3.85	35.87	580.29	169.63	17930.50	1551.03	20271.17	79%	782.95	16%

- Ak8 jet passing selection w/ max(PNet Hbb vs. QCD) is selected as the Hbb jet
- Variables left to play with (next slides):
  - VBS kinematics: M_{jj}, $\Delta\eta_{jj}$
  - WH kinematics: $L_T = p_T(\ell)$ + MET, $S_T = L_T + p_T$(Hbb fat jet), Hbb vs. QCD ParticleNet score

*eff = 1 - after/before
**ΔR < 0.8 lepton overlap removal
***ΔR < 0.4 lepton overlap removal, ΔR < 0.8 Hbb fat jet overlap removal

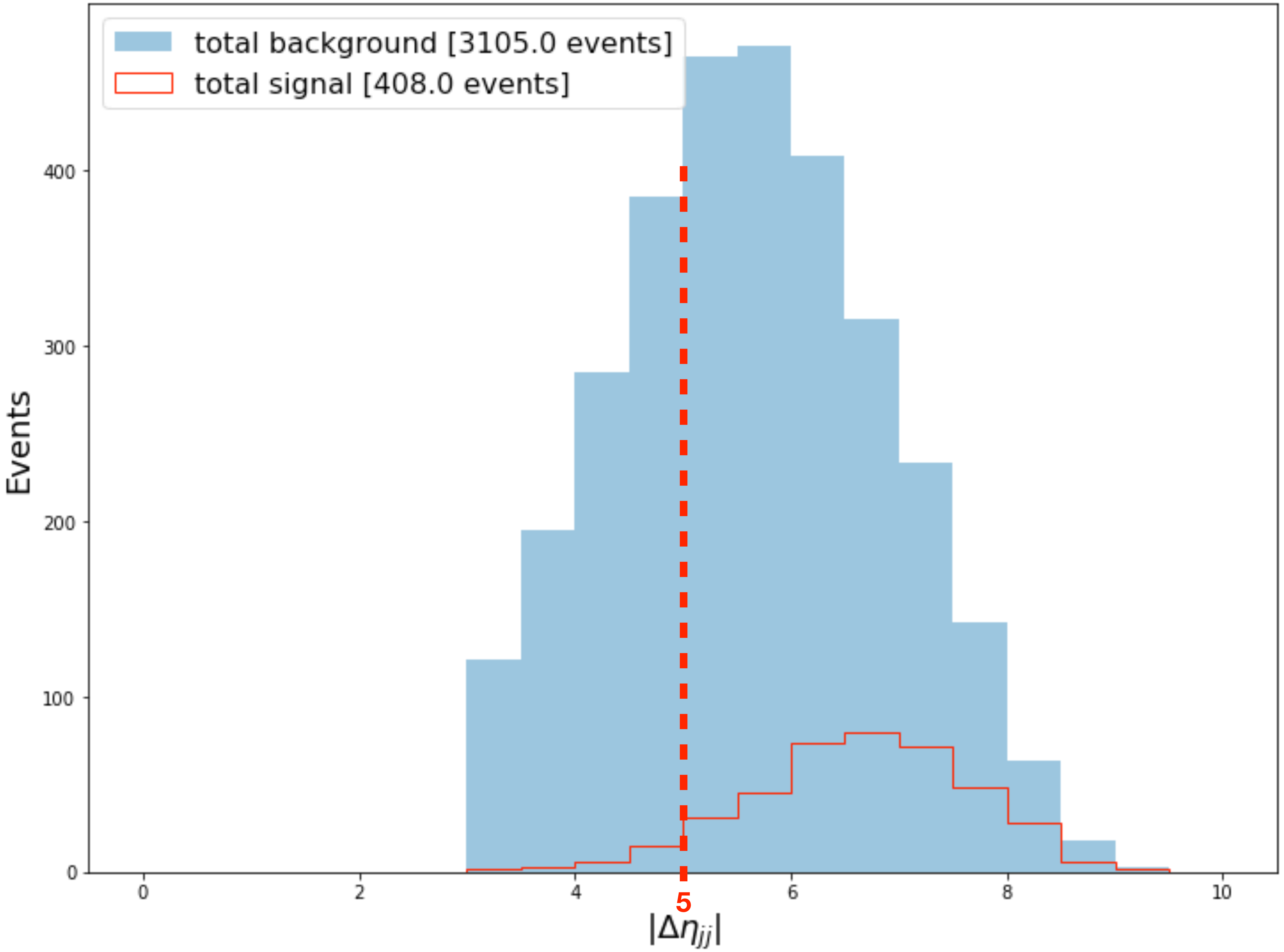
UC San Diego

VBS Kinematics: M_{jj}



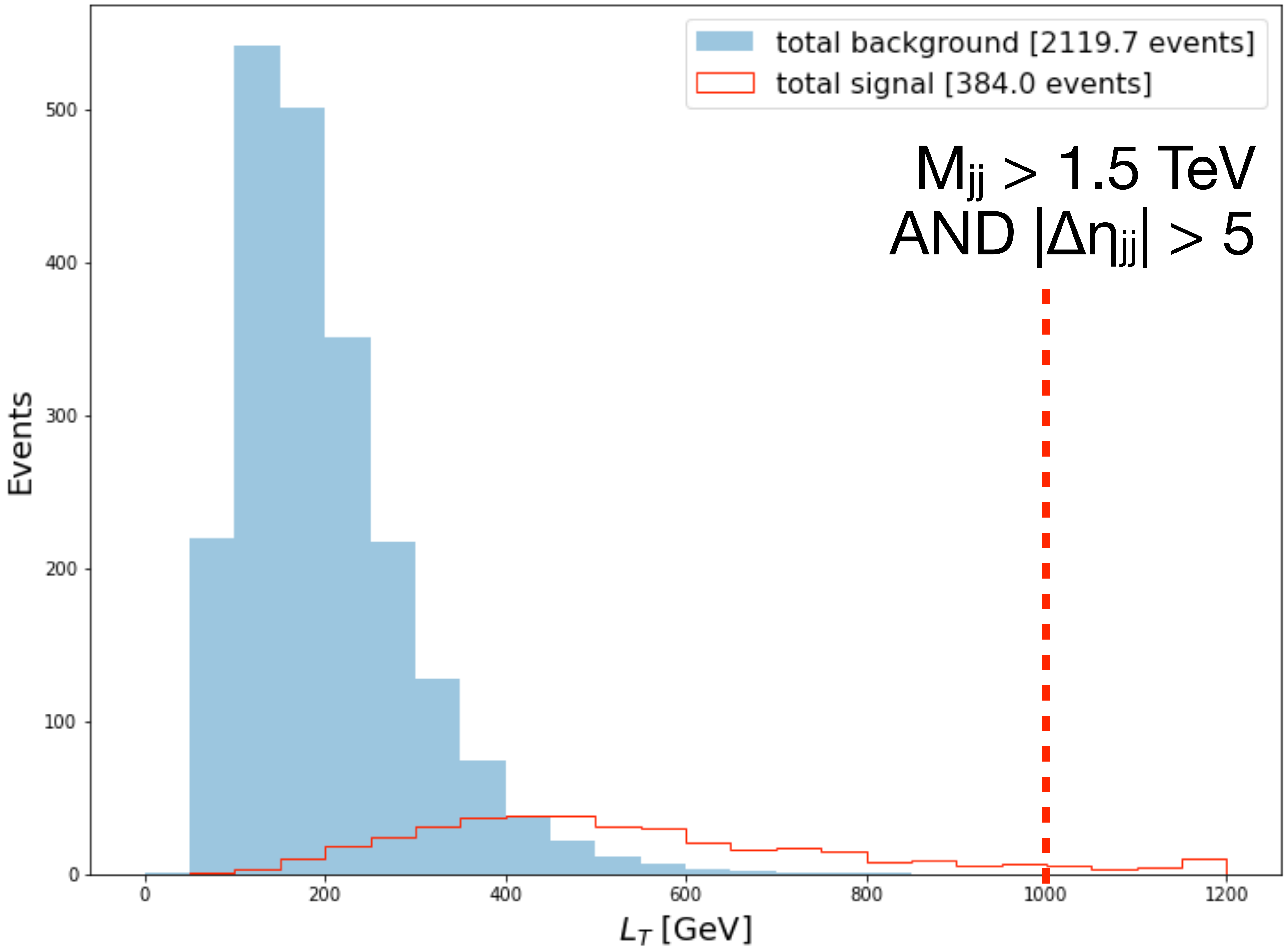
Cut	VBSWH (kW < 0)	TotalBkg	S/√B
Preselection	782.95	20271.17	5.50
M_{jj} > 1 TeV	606.26	7731.96	6.89
M_{jj} > 1.5 TeV	408.00	3104.98	7.32
M_{jj} > 2 TeV	244.29	1330.55	6.70
M_{jj} > 2.5 TeV	134.39	589.18	5.54

VBS Kinematics: $|\Delta\eta_{jj}|$



Cut	VBSWH (kW < 0)	TotalBkg	S/√B
M_{jj} > 1.5 TeV	408.00	3104.98	7.32
$\lvert\Delta\eta_{jj}\rvert$ > 3.5	406.47	2984.11	7.44
$\lvert\Delta\eta_{jj}\rvert$ > 4	403.92	2789.56	7.65
$\lvert\Delta\eta_{jj}\rvert$ > 4.5	398.54	2504.95	7.96
$\lvert\Delta\eta_{jj}\rvert$ > 5	384.00	2119.74	8.34
$\lvert\Delta\eta_{jj}\rvert$ > 5.5	352.89	1654.81	8.67
$\lvert\Delta\eta_{jj}\rvert$ > 6	307.70	1183.25	8.95
$\lvert\Delta\eta_{jj}\rvert$ > 6.5	234.21	775.42	8.41
$\lvert\Delta\eta_{jj}\rvert$ > 7	154.36	460.01	7.20

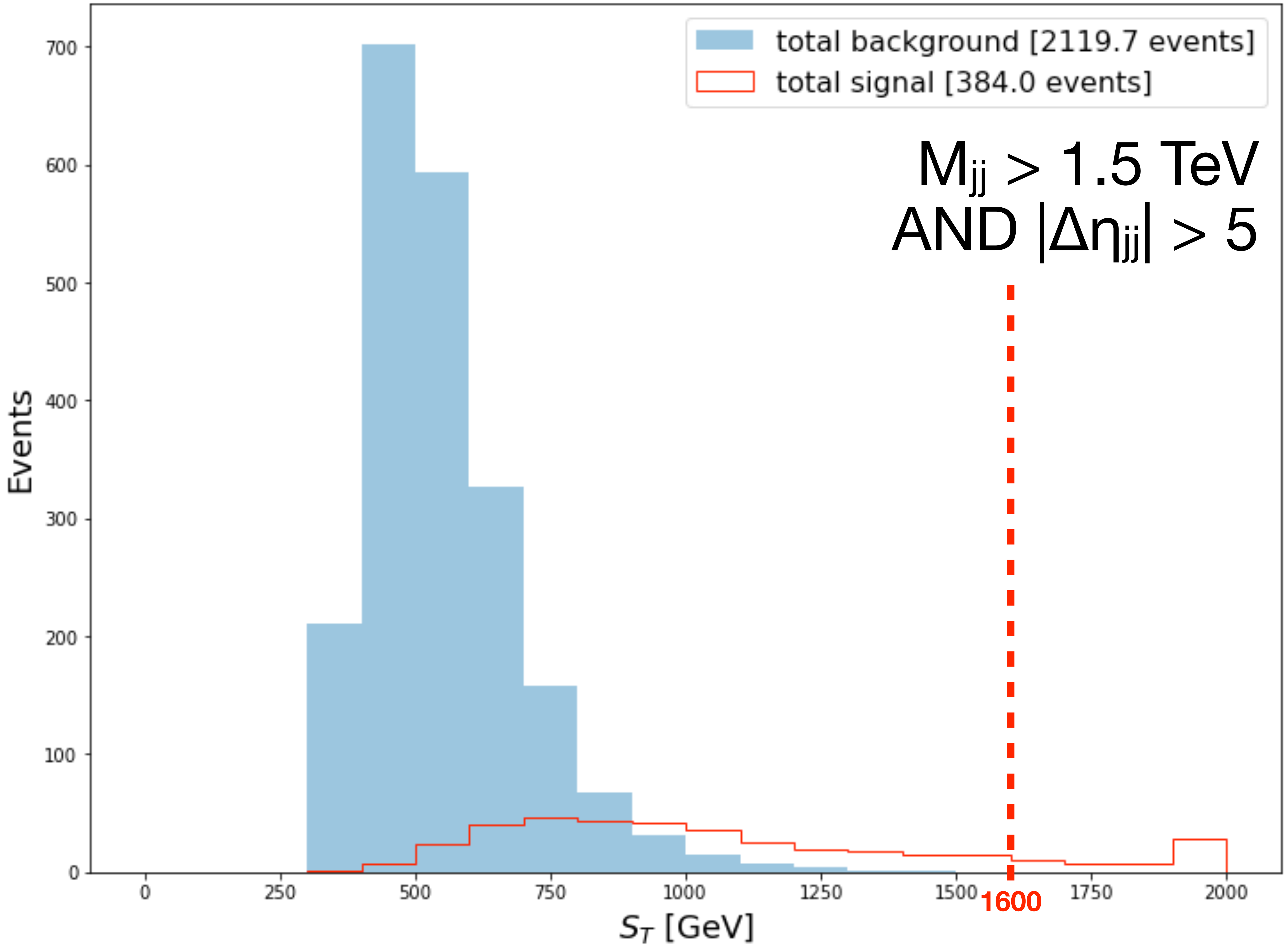
WH Kinematics: $L_T = p_T(\ell) + MET$



Cut	VBSWH (kW < 0)	TotalBkg	S/√B
M_{jj} > 1.5 TeV AND $\Delta\eta_{jj}$ > 5	384.00	2119.74	8.34
L_T > 500 GeV	181.81	25.95	35.69
L_T > 600 GeV	121.16	7.91	43.07
L_T > 700 GeV	85.36	2.57	53.28
L_T > 800 GeV	53.70	0.96	54.74
L_T > 900 GeV	36.30	0.33	63.08
L_T > 1 TeV	23.54	0.12	67.60
L_T > 1.1 TeV	15.26	0.11	45.35
L_T > 1.2 TeV	8.63	0.07	32.03

M_{jj} > 1.5 TeV AND $\Delta\eta_{jj}$ > 5
AND L_T > 900 GeV
gives **36 sig vs. ~0 bkg**

WH Kinematics: $S_T = L_T + p_T$(Hbb Ak8 jet)



Cut	VBSWH (kW < 0)	TotalBkg	S/√B
M_{jj} > 1.5 TeV AND $\Delta\eta_{jj}$ > 5	384.00	2119.74	8.34
S_T > 1 TeV	180.34	28.92	33.54
S_T > 1.1 TeV	144.59	13.81	38.91
S_T > 1.2 TeV	119.09	6.94	45.20
S_T > 1.3 TeV	99.89	3.33	54.76
S_T > 1.4 TeV	82.31	1.41	69.40
S_T > 1.5 TeV	67.60	0.65	83.86
S_T > 1.6 TeV	53.34	0.28	101.20
S_T > 1.7 TeV	43.10	0.07	167.06
S_T > 1.8 TeV	36.05	0.11	111.07
S_T > 1.9 TeV	28.40	0.10	89.40

M_{jj} > 1.5 TeV AND $\Delta\eta_{jj}$ > 5
AND S_T > 1.6 TeV
gives **53 sig vs. ~0 bkg**

CMS

UC San Diego

VBS WH BDT Preselection

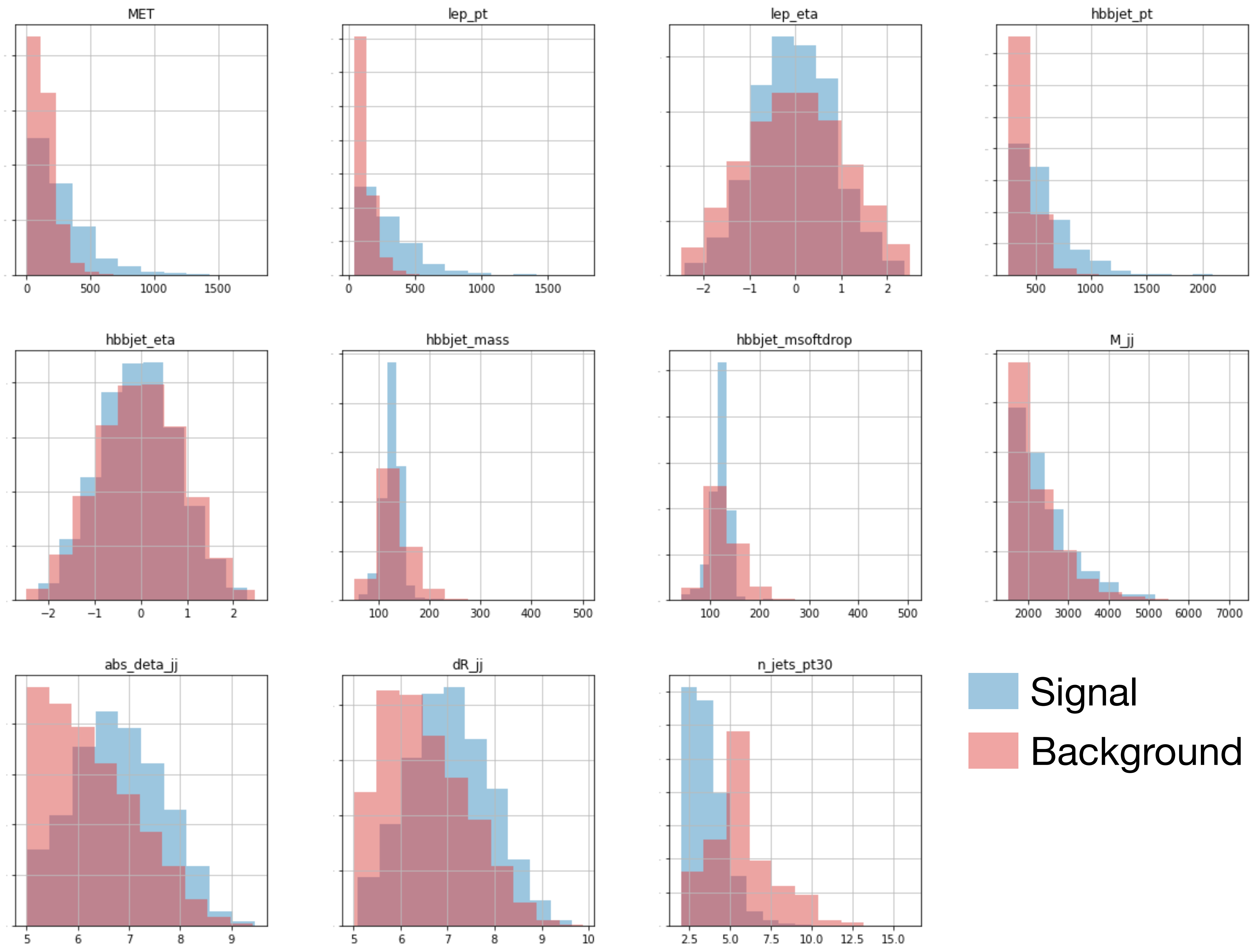
Cut	VHToNonbb	VV/VVV/VBSWZ	W+Jets	TTbar+X	TTbar1L	TTbar2L	TotalBkg	Eff.*	VBSWH (kW < 0)	Eff.*
Skim AND Postskim	167.20	1682.81	32660.21	3675.78	424281.50	34177.24	496644.74	—	59598.09	—
Preselection	3.85	35.87	580.29	169.63	17930.50	1551.03	20271.17	96%	782.95	99%
M_{jj} > 1.5 TeV AND $\Delta\eta_{jj}$ > 5 AND S_T > 500 GeV	0.00	1.80	30.64	10.19	1058.81	105.57	1207.01	94%	376.27	52%

- Trained a simple XGBoost BDT to distinguish signal from background
- Split MC into training (60%) and testing (40%) samples
  - Rescale integrals sample-by-sample to match original integral
    - i.e. if Σ_{orig}SampleX = 100.5 weighted events, then rescale event weights such that Σ_{train}SampleX = Σ_{test}SampleX = 100.5
  - Added tighter VBS and S_T cuts after Preselection

*eff = 1 - after/before

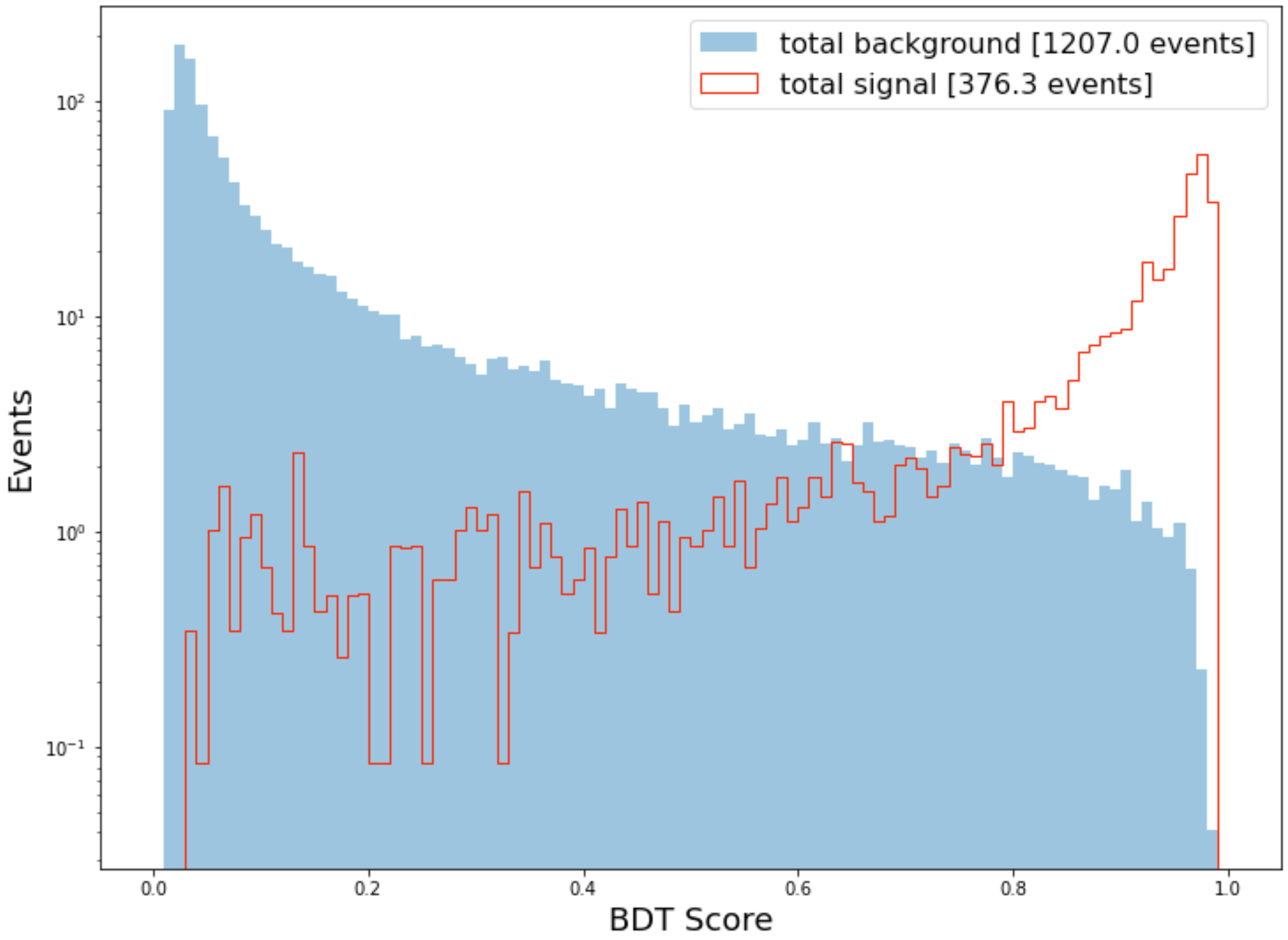
VBS WH BDT: Features

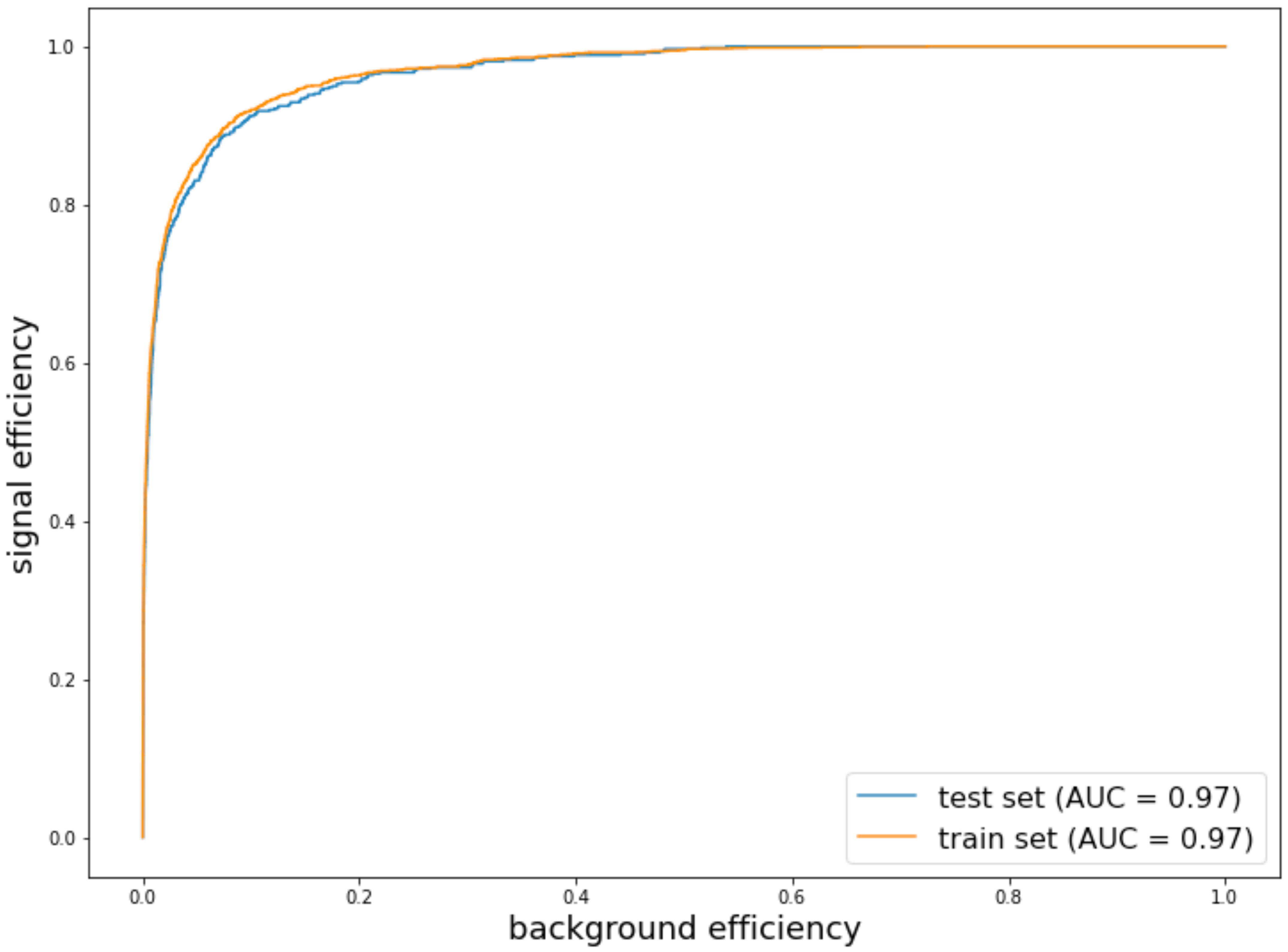
Features	
Object	**Variables**
Lepton	• p_T • η
Hbb Ak8 Jet	• p_T • η • mass • M_{SD}
VBS jets	• M_{jj} • $\|\Delta\eta_{jj}\|$ • ΔR_{jj}
Other	• Number of jets w/ p_T > 30 GeV • MET



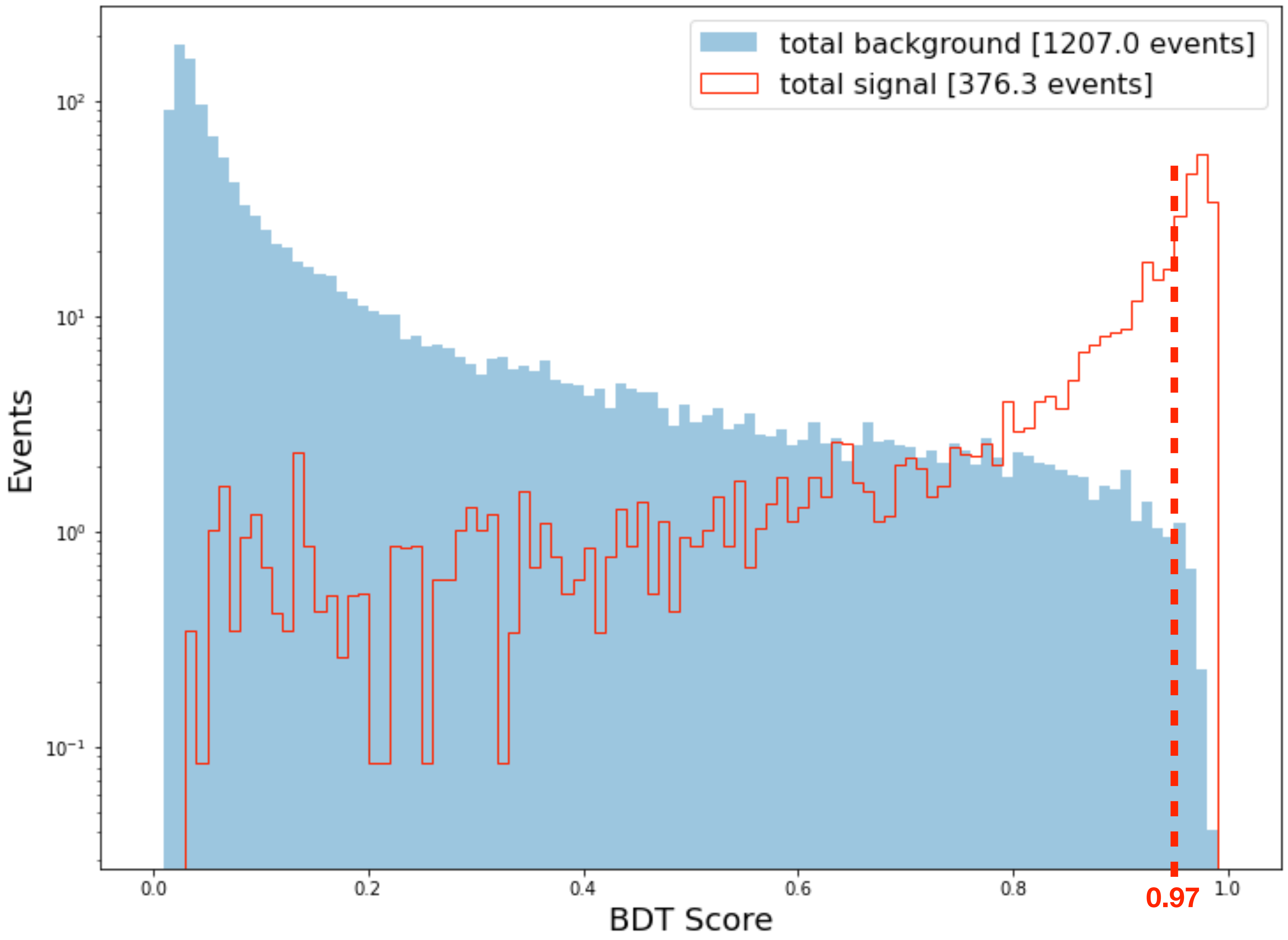
UC San Diego

VBS WH BDT: Results





VBS WH BDT: Results



Cut	VBSWH (kW < 0)	TotalBkg	S/√B
BDT Preselection	376.27	1207.01	10.83
BDT > 0.80	289.30	27.22	55.45
BDT > 0.90	235.71	8.43	81.19
BDT > 0.91	227.04	6.50	89.04
BDT > 0.92	215.28	5.39	92.76
BDT > 0.93	197.41	4.01	98.60
BDT > 0.94	182.57	2.97	105.92
BDT > 0.95	166.04	2.03	116.52
BDT > 0.96	136.64	0.93	141.36
BDT > 0.97	90.71	0.27	175.18
BDT > 0.98	33.87	0.04	166.49

BDT > 0.97 gives **91 sig vs. ~0 bkg**

Summary

- Arrived at BDT- or cut-based signal regions that should be able to exclude $\kappa_W < 0$
- Immediate next steps:
  - Apply scale factors: lepton ID (ttH), JECs, etc.
  - Freeze signal region optimization
  - Design control region(s) for background estimation
- Afterwards:
  - Make another presentation here
  - Ask for a CADI line

Backup

*Professor
†Post-doc
‡Graduate Student
§Undergraduate Student

Who is involved?

- UC San Diego:
  A. Arora‡, P. Chang†, L. Giannini†, J. Guiang‡, F. Würthwein*, Y. Xiang‡, A. Yagil*

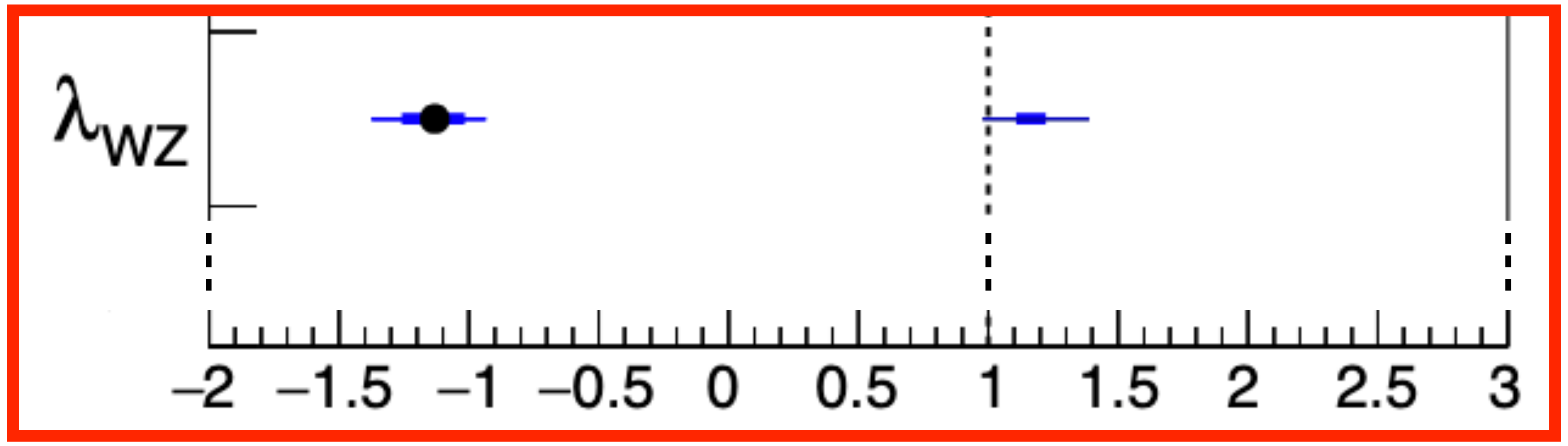
- U. Nebraska:
  F. Golf*

- Boston University:
  D. Spitzbart†, I. Suarez*

- UC Santa Barbara:
  C. Campagnari*

UC San Diego Nebraska (University of Nebraska Lincoln) BOSTON UNIVERSITY UC SANTA BARBARA

Current CMS Result



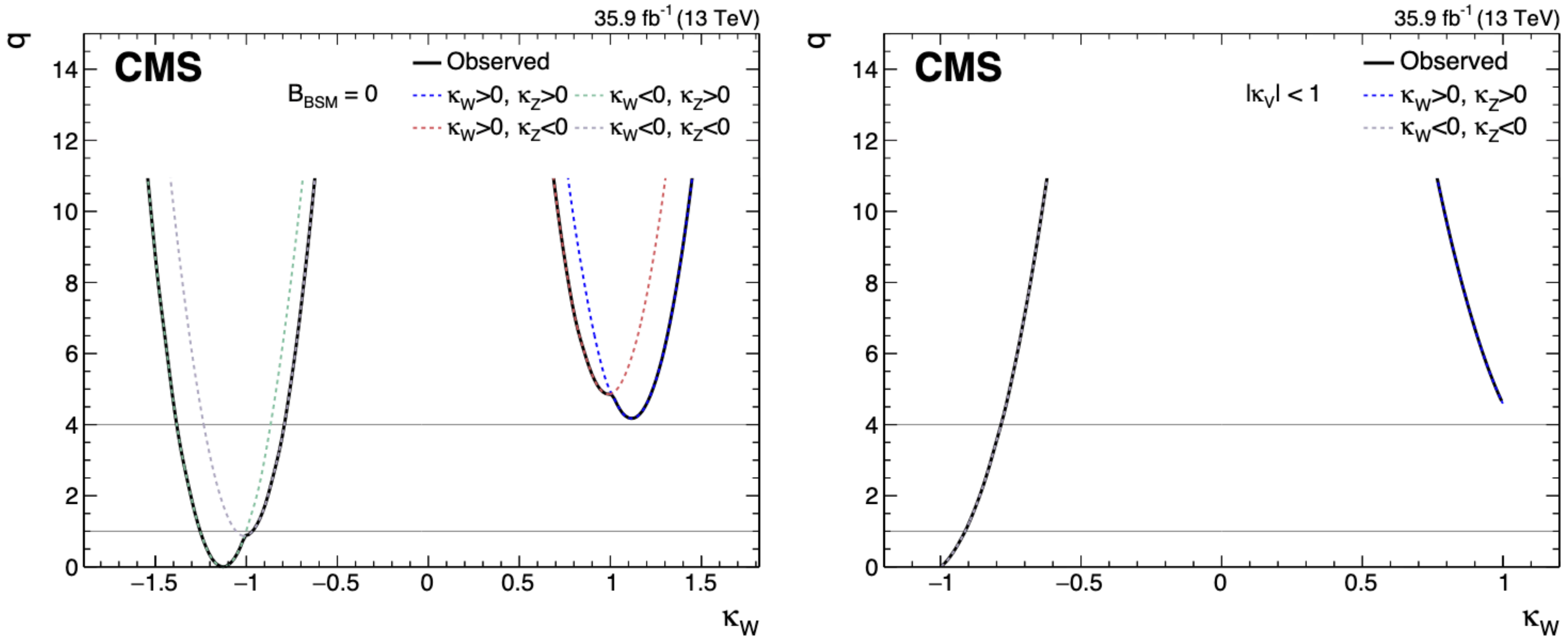


Figure 13: Scan of the test statistic q as a function of κ_W in the generic κ model assuming $\mathcal{B}_{BSM} = 0$ (left) and allowing $\mathcal{B}_{inv}$ and $\mathcal{B}_{undet}$ to float (right). The different colored lines indicate the value of q for different combinations of signs for κ_W and κ_Z. The solid black line shows the minimum value of $q(\kappa_W)$ in each case and is used to determine the best fit point and the 1σ and 2σ CL regions. The scan in the right figure is truncated because of the constraints of $|\kappa_W| \leq 1$ and $|\kappa_Z| \leq 1$, which are imposed in this model.

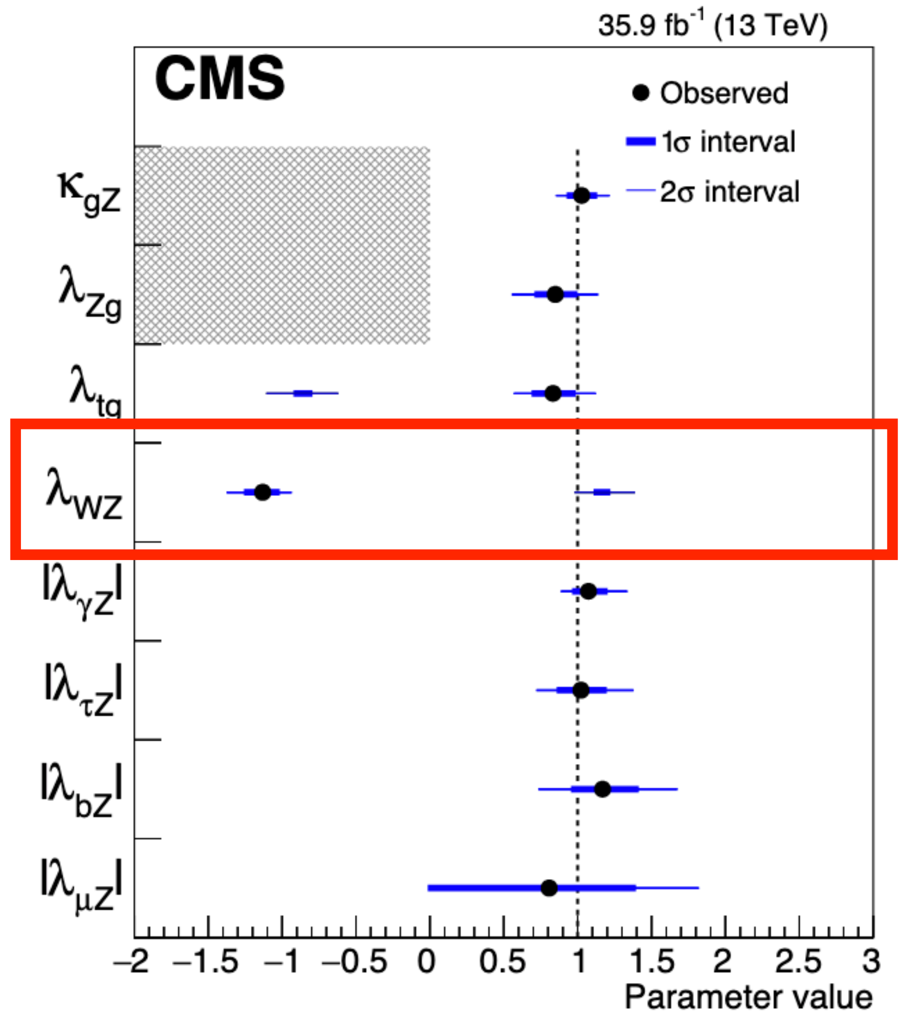


Figure 16: Summary of the model with coupling ratios and effective couplings for the ggH and $H \to \gamma\gamma$ loops. The points indicate the best fit values while the thick and thin horizontal bars show the 1σ and 2σ CL intervals, respectively. For this model, both positive and negative values of λ_{WZ} and λ_{tg} are considered.

https://arxiv.org/pdf/1809.10733.pdf

BSM Signal Models

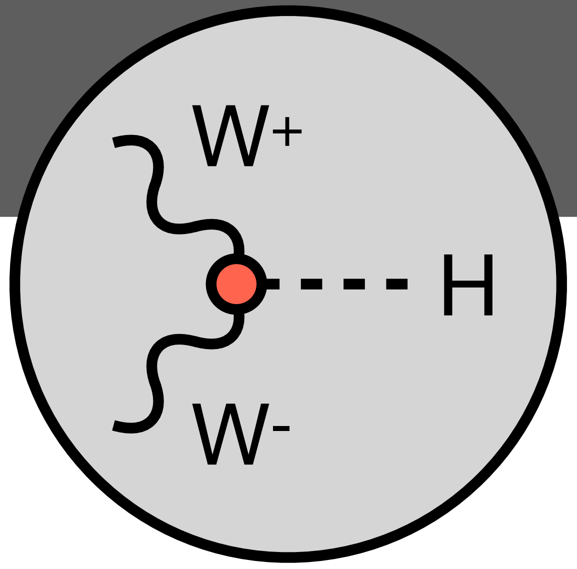
$\kappa_W < 0$

models/sm/couplings.py

```
GC_72 = Coupling(name = 'GC_72',
                 value = '(ee**2*complex(0,1)*vev)/(2.*sw**2)',
                 value = '-((ee**2*complex(0,1)*vev)/(2.*sw**2))',
                 order = {'QED':1})
```

models/sm/vertices.py

```
V_52 = Vertex(name = 'V_52',
              particles = [ P.W__minus__, P.W__plus__, P.H ],
              color = [ '1' ],
              lorentz = [ L.VVS1 ],
              couplings = {(0,0):C.GC_72})
```



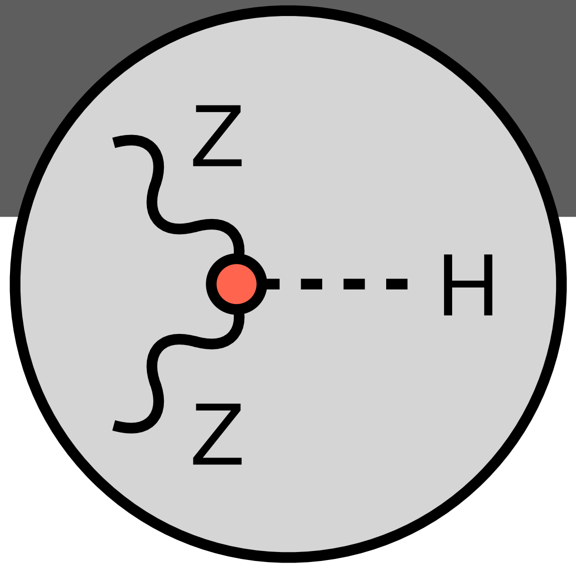
$\kappa_Z < 0$

models/sm/couplings.py

```
GC_81 = Coupling(name = 'GC_81',
                 value = 'ee**2*complex(0,1)*vev + …',
                 value = '-(ee**2*complex(0,1)*vev + … )',
                 order = {'QED':1})
```

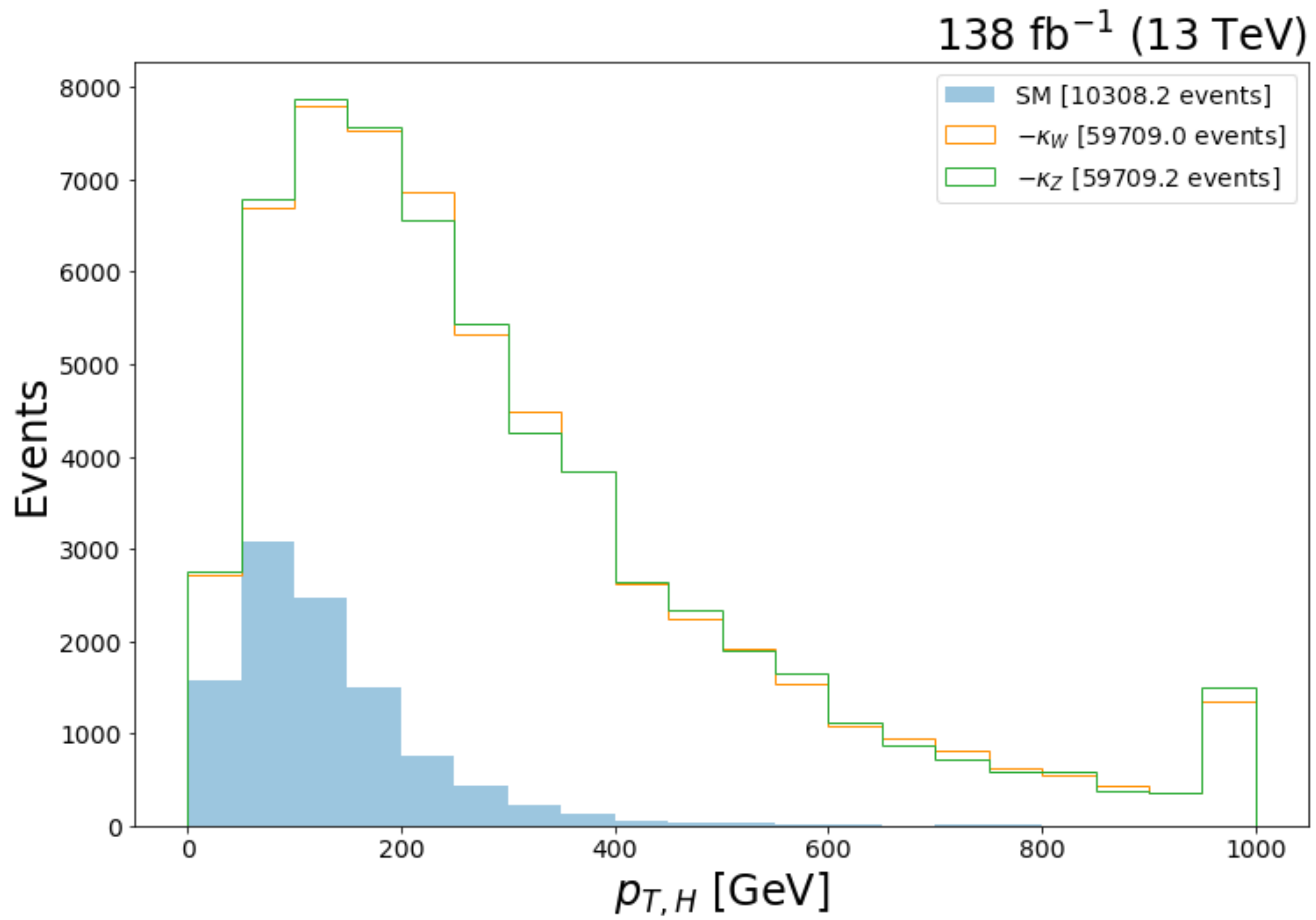
models/sm/vertices.py

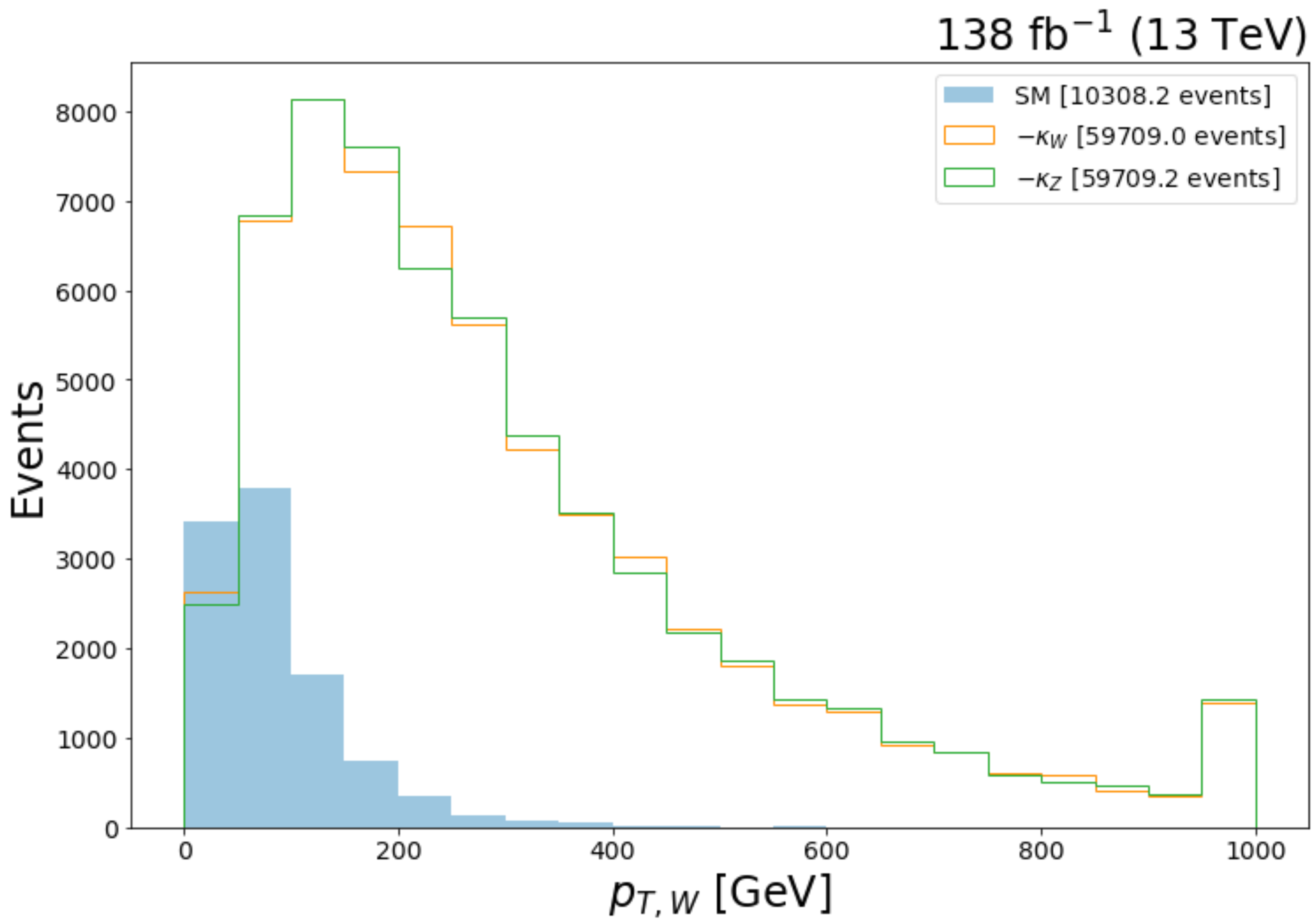
```
V_69 = Vertex(name = 'V_69',
              particles = [ P.Z, P.Z, P.H ],
              color = [ '1' ],
              lorentz = [ L.VVS1 ],
              couplings = {(0,0):C.GC_81})
```



Only changed one line in SM Madgraph model!

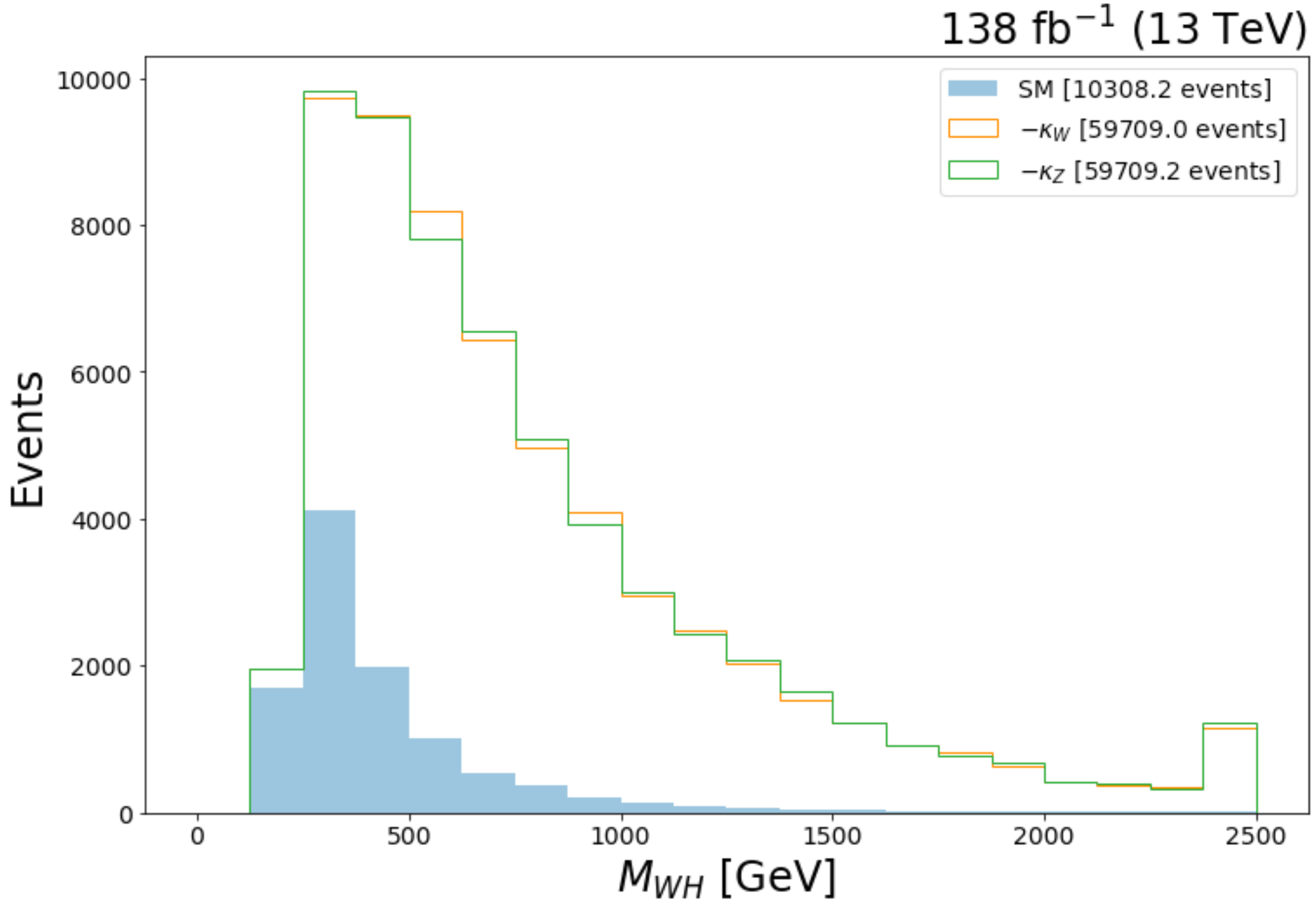
LHE W and H Kinematics





Strong boost in W and H in BSM ($-\kappa_W$ or $-\kappa_Z$) models

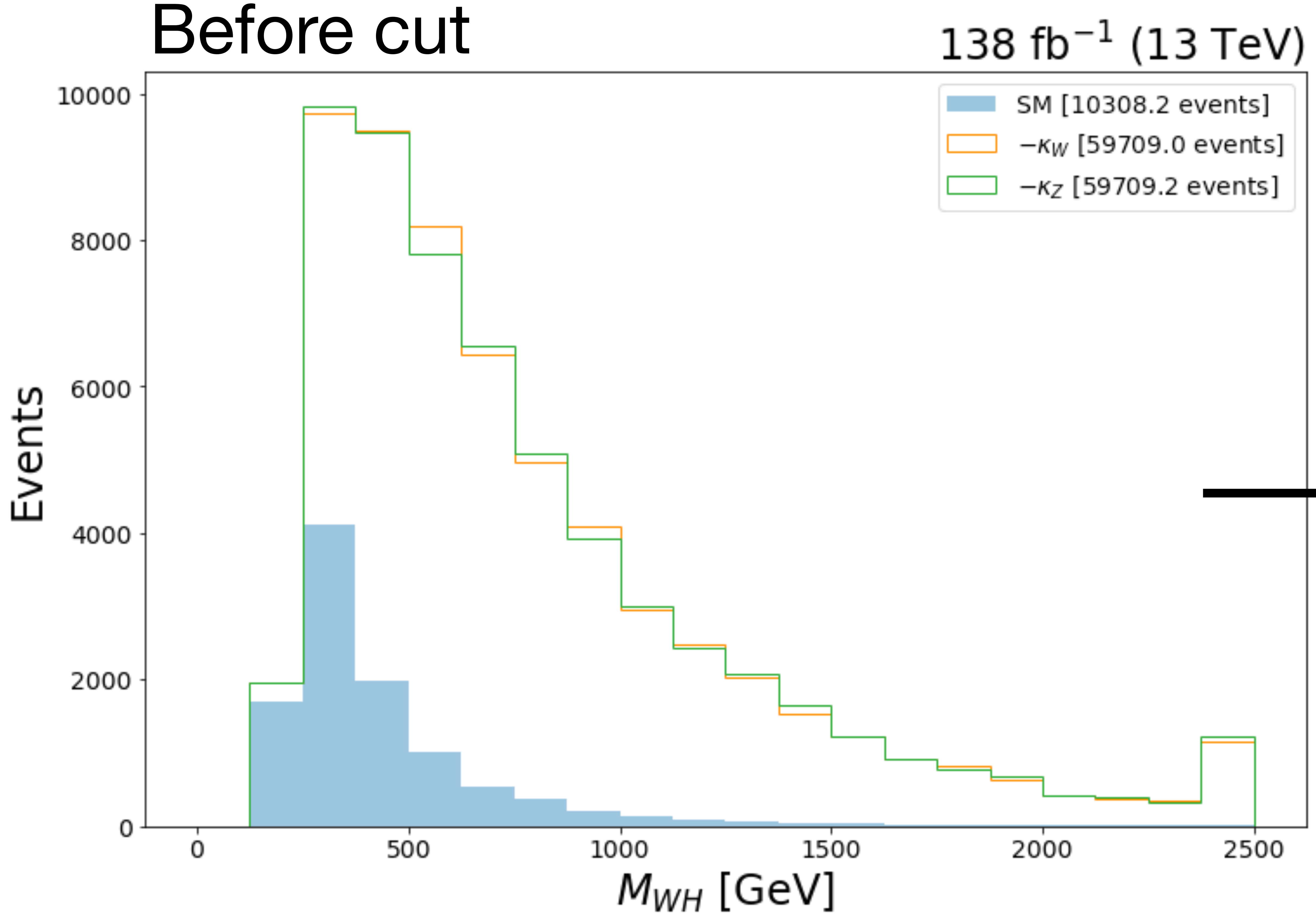
LHE W and H Kinematics



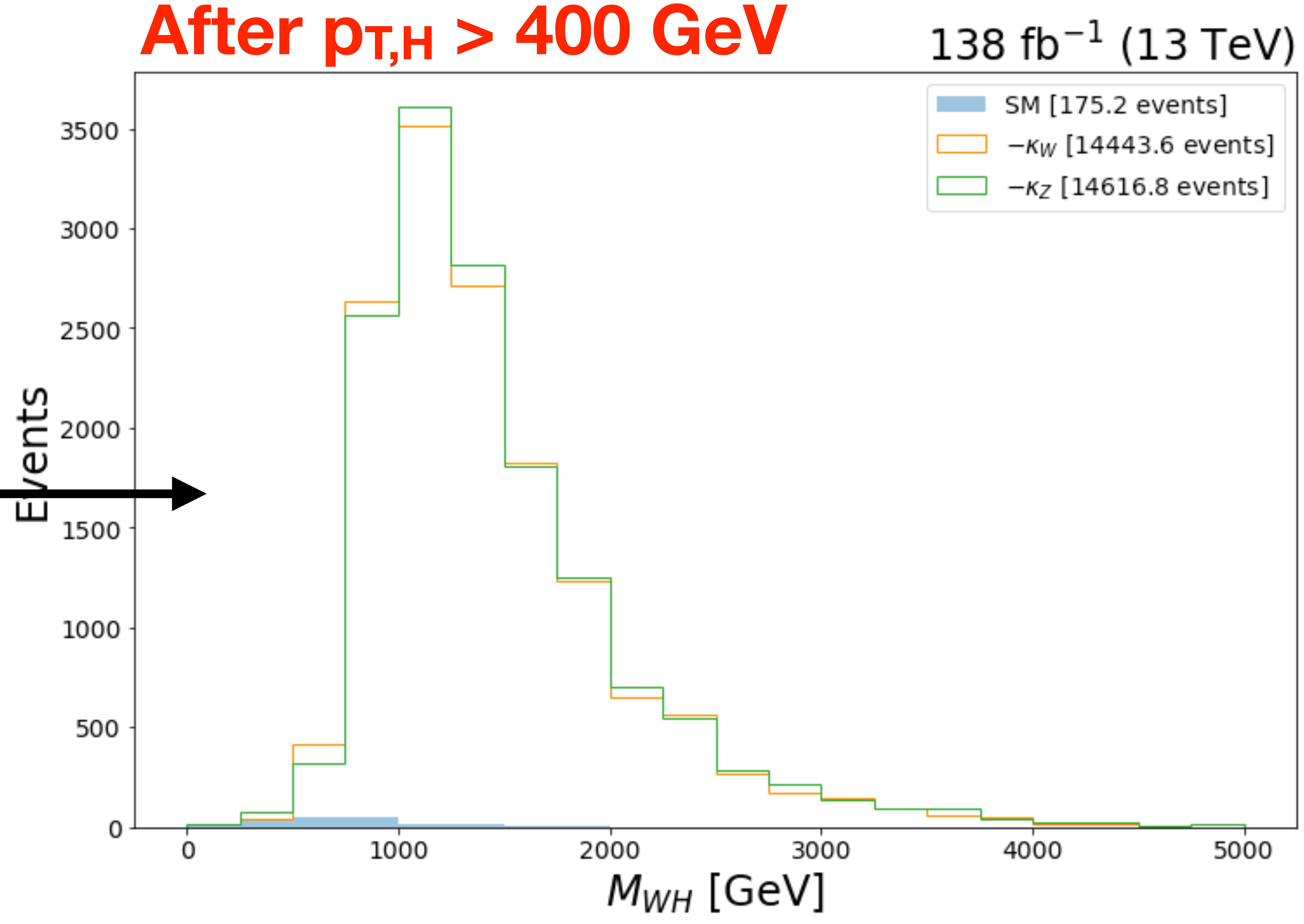
Strong boost in W and H in BSM ($-\kappa_W$ or $-\kappa_Z$) models

LHE W and H Kinematics

Before cut



After $p_{T,H}$ > 400 GeV

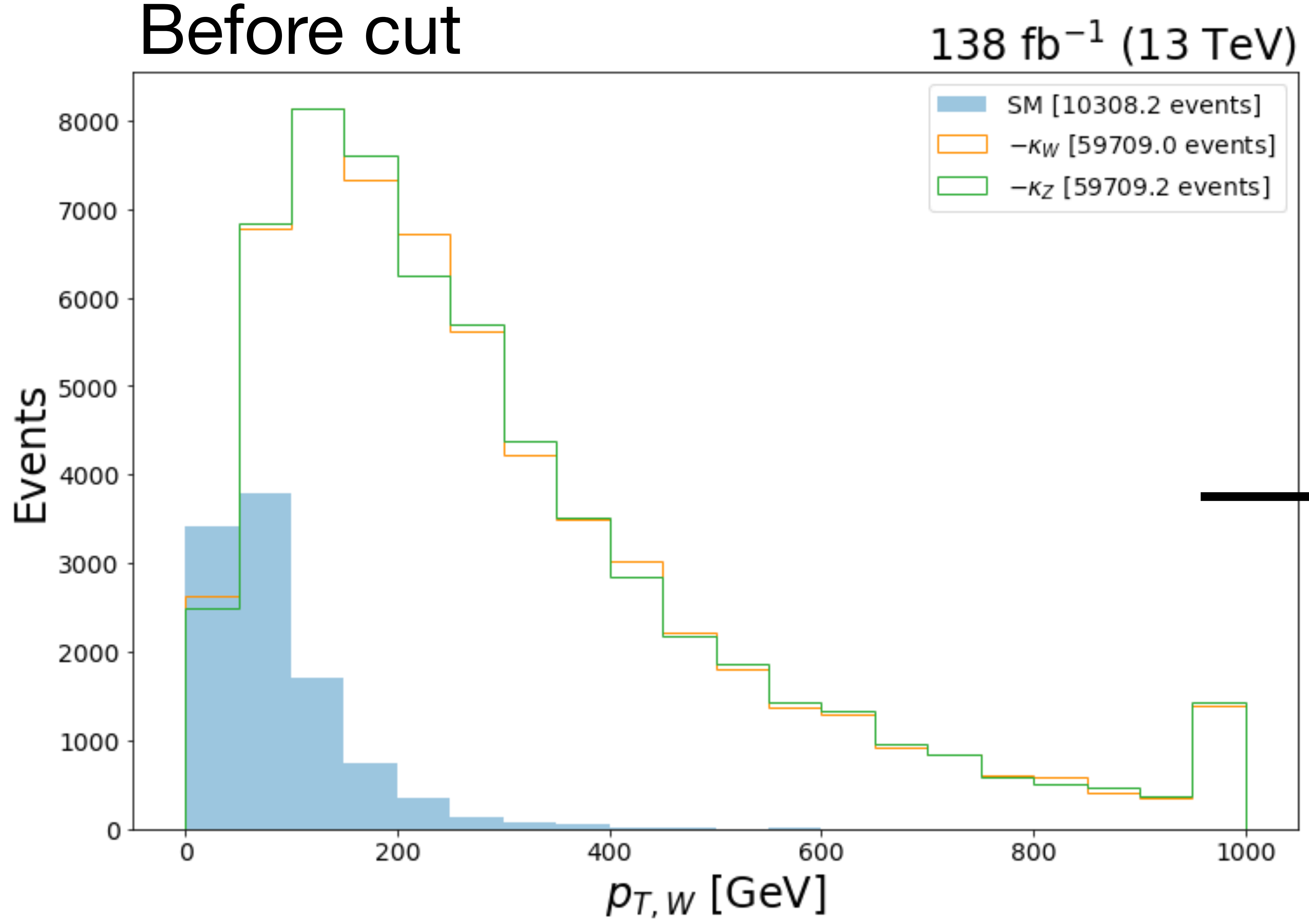


High s-hat events have large M_{WH}

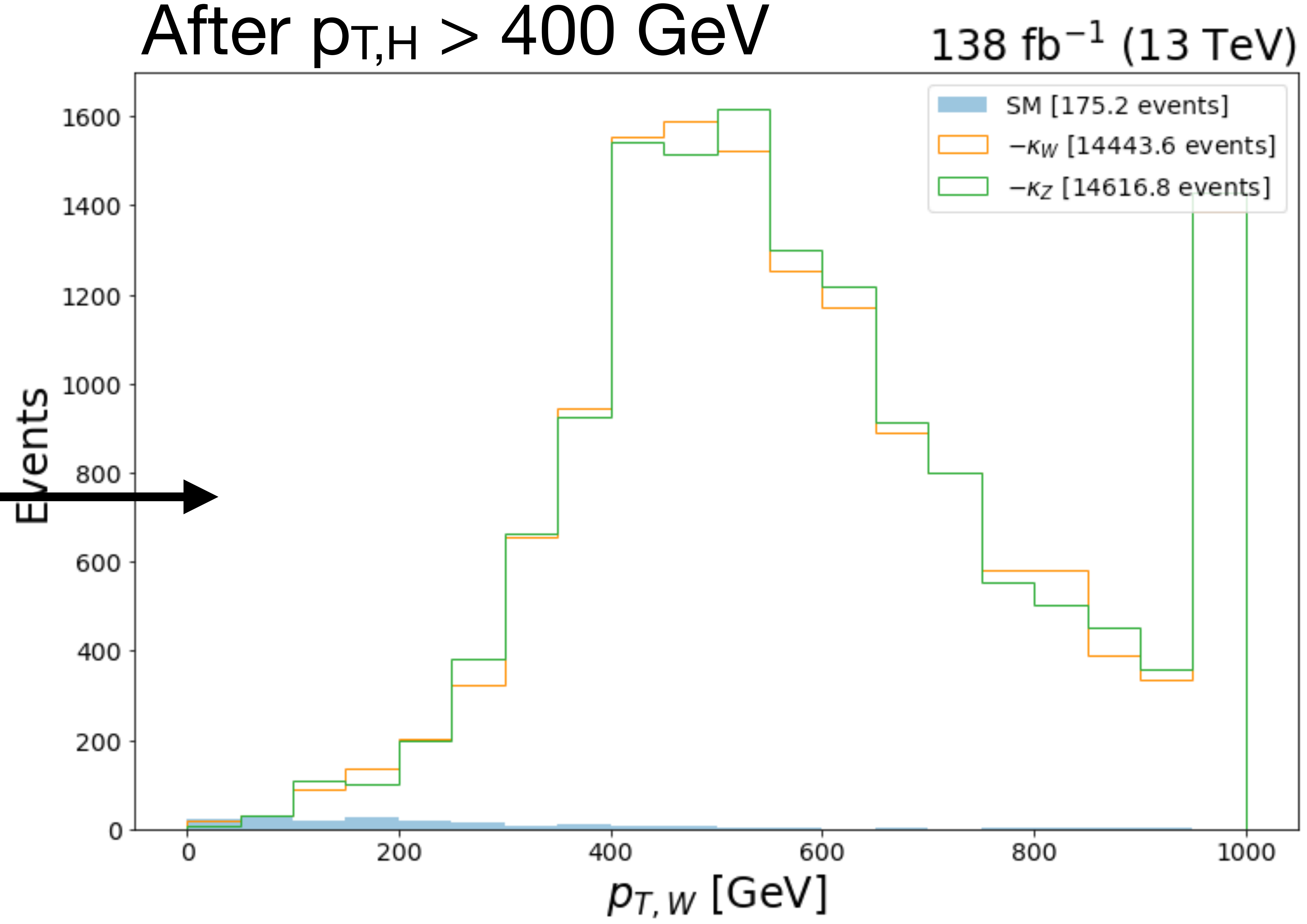
UC San Diego

LHE W and H Kinematics

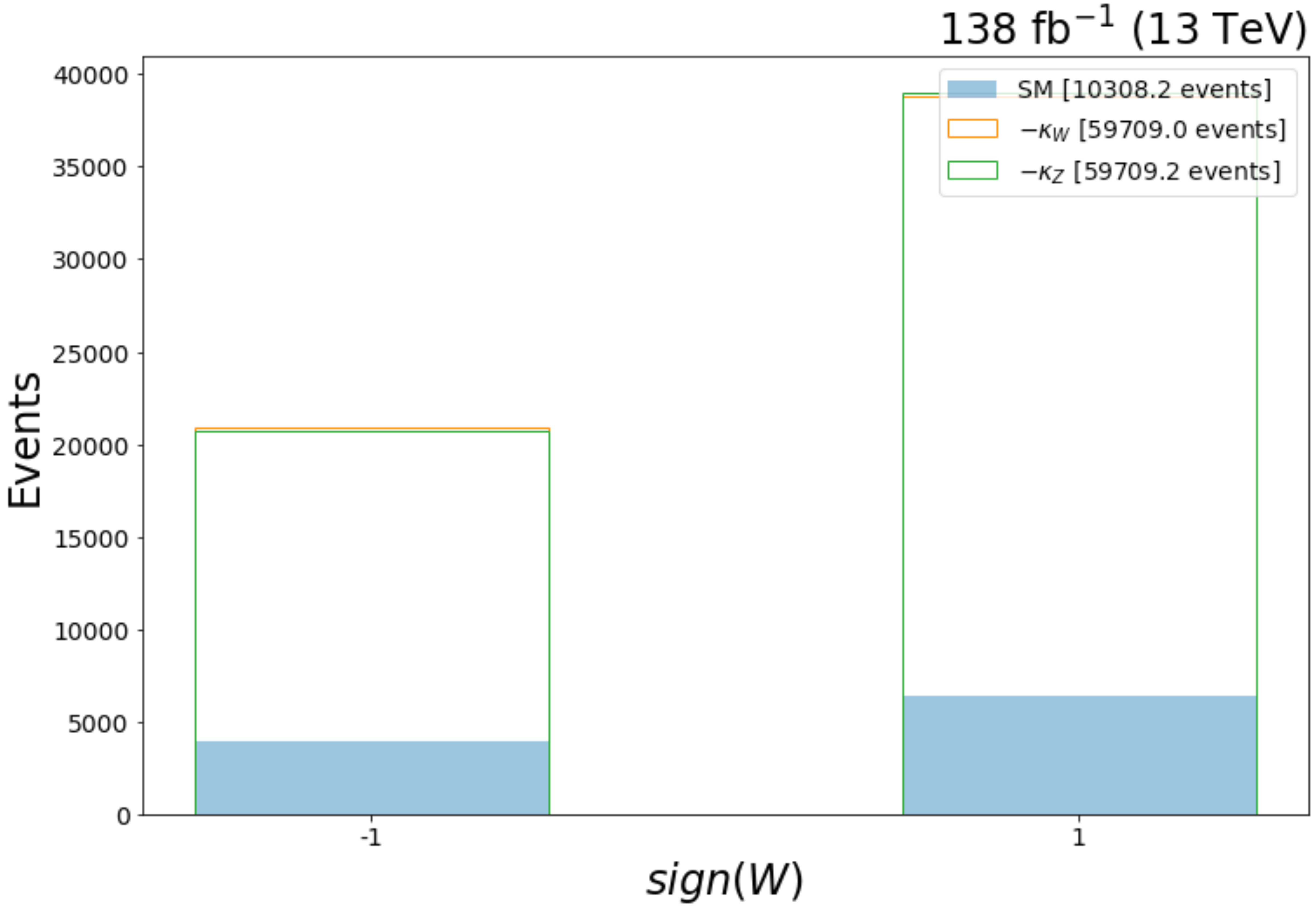
Before cut

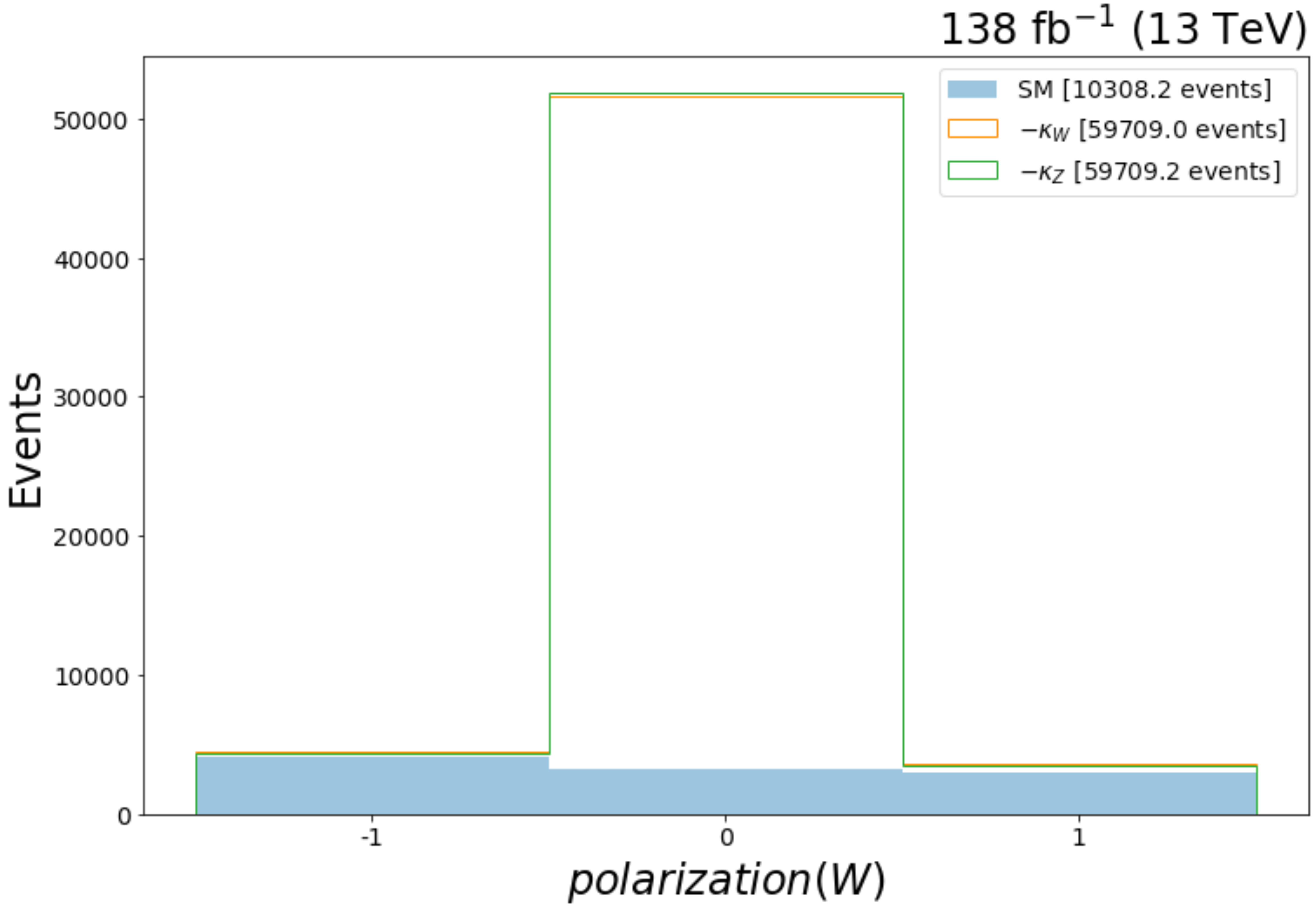


After $p_{T,H}$ > 400 GeV



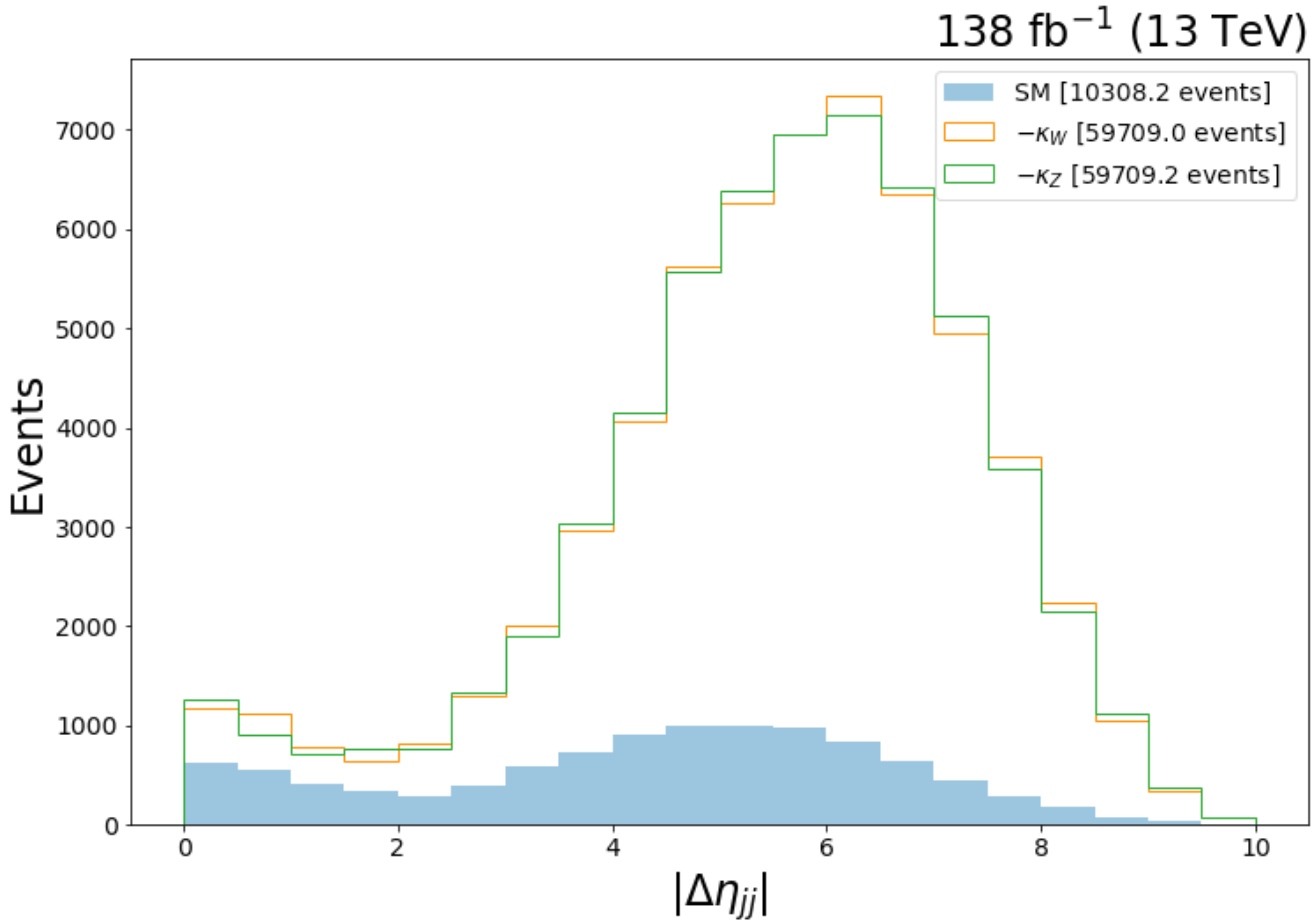
LHE W Characteristics

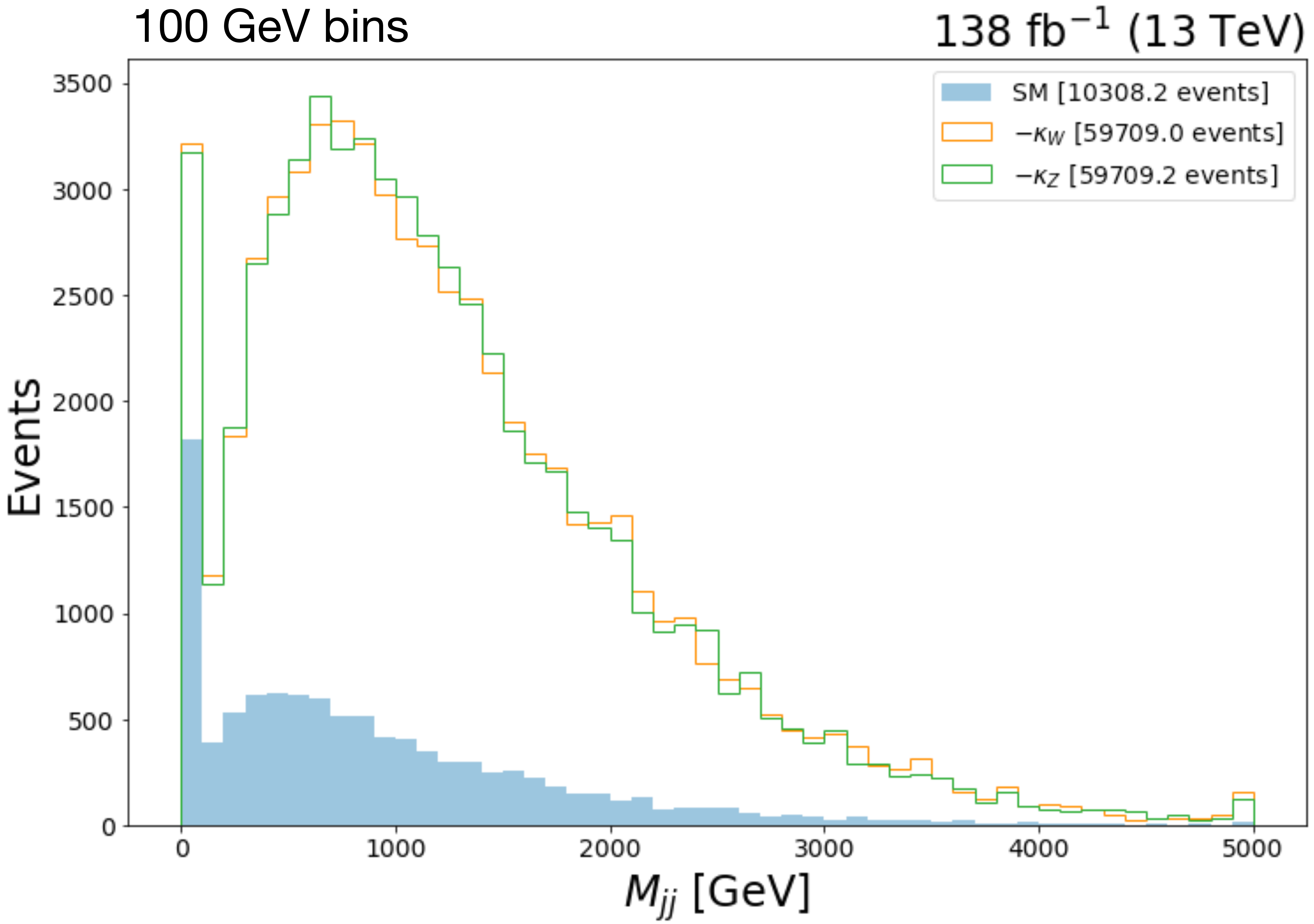




Some W+/W- asymmetry; more longitudinally polarized W's

LHE VBS Kinematics

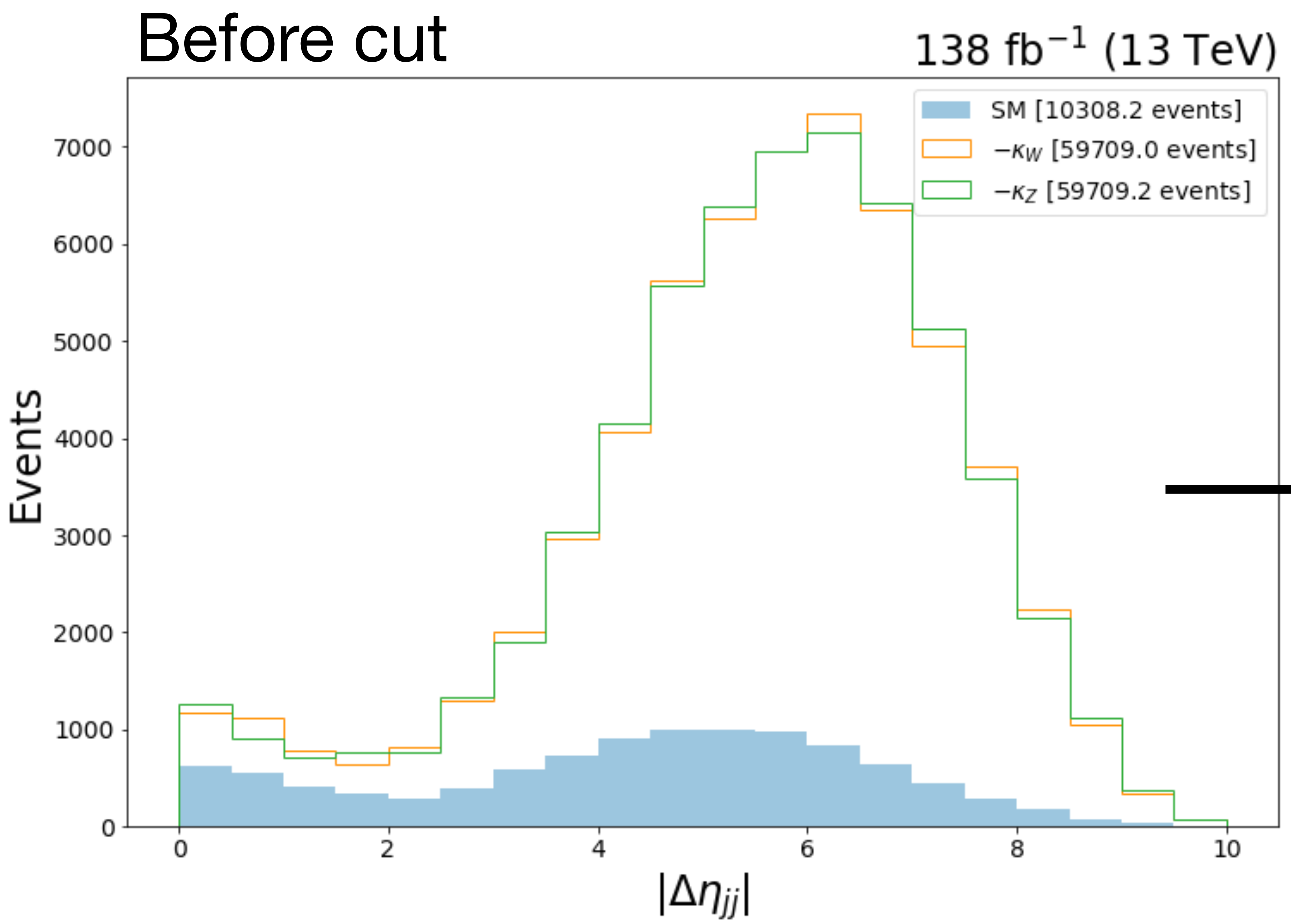


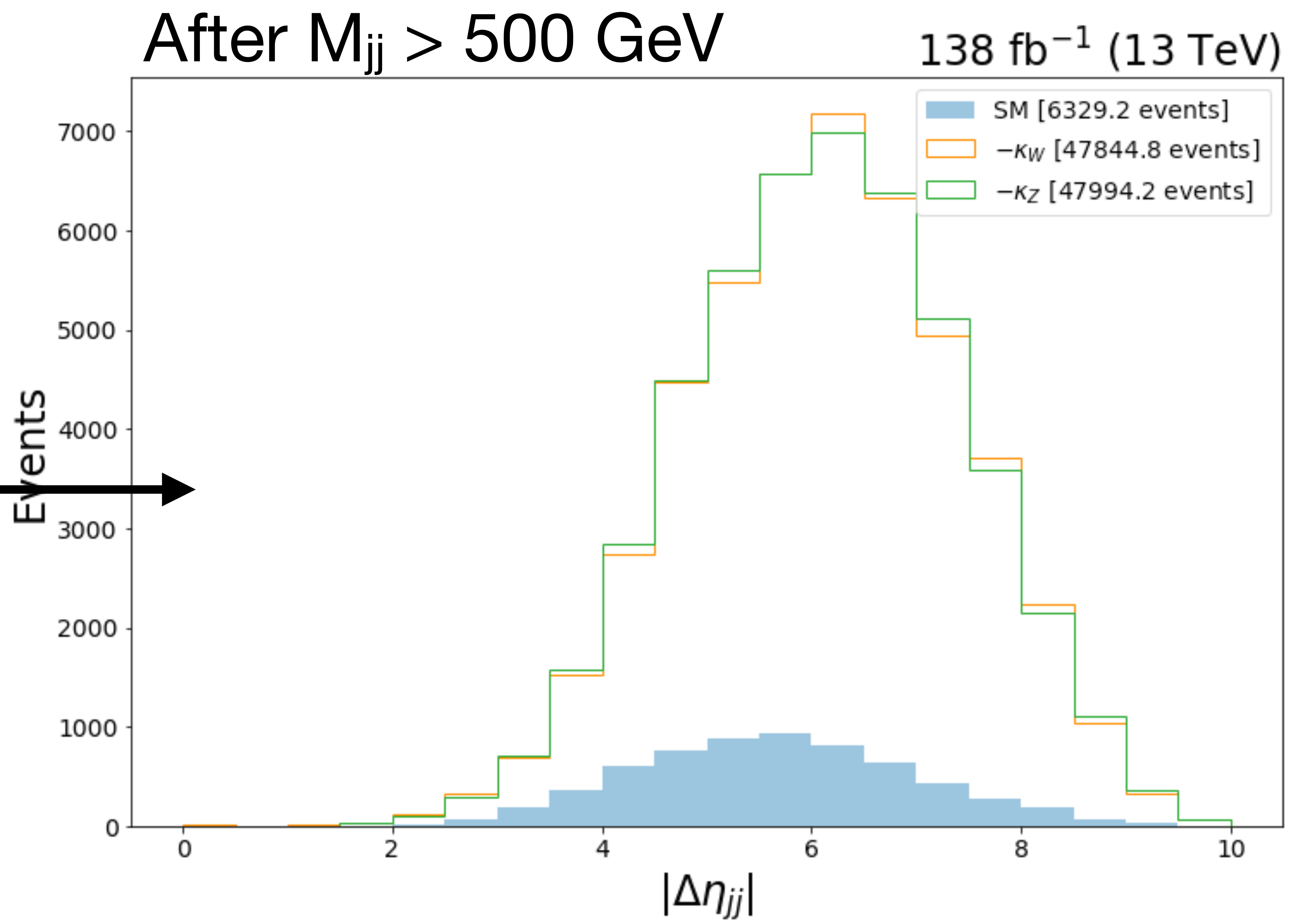


VBS-like signal: large $|\Delta\eta_{jj}|$, M_{jj}

UC San Diego

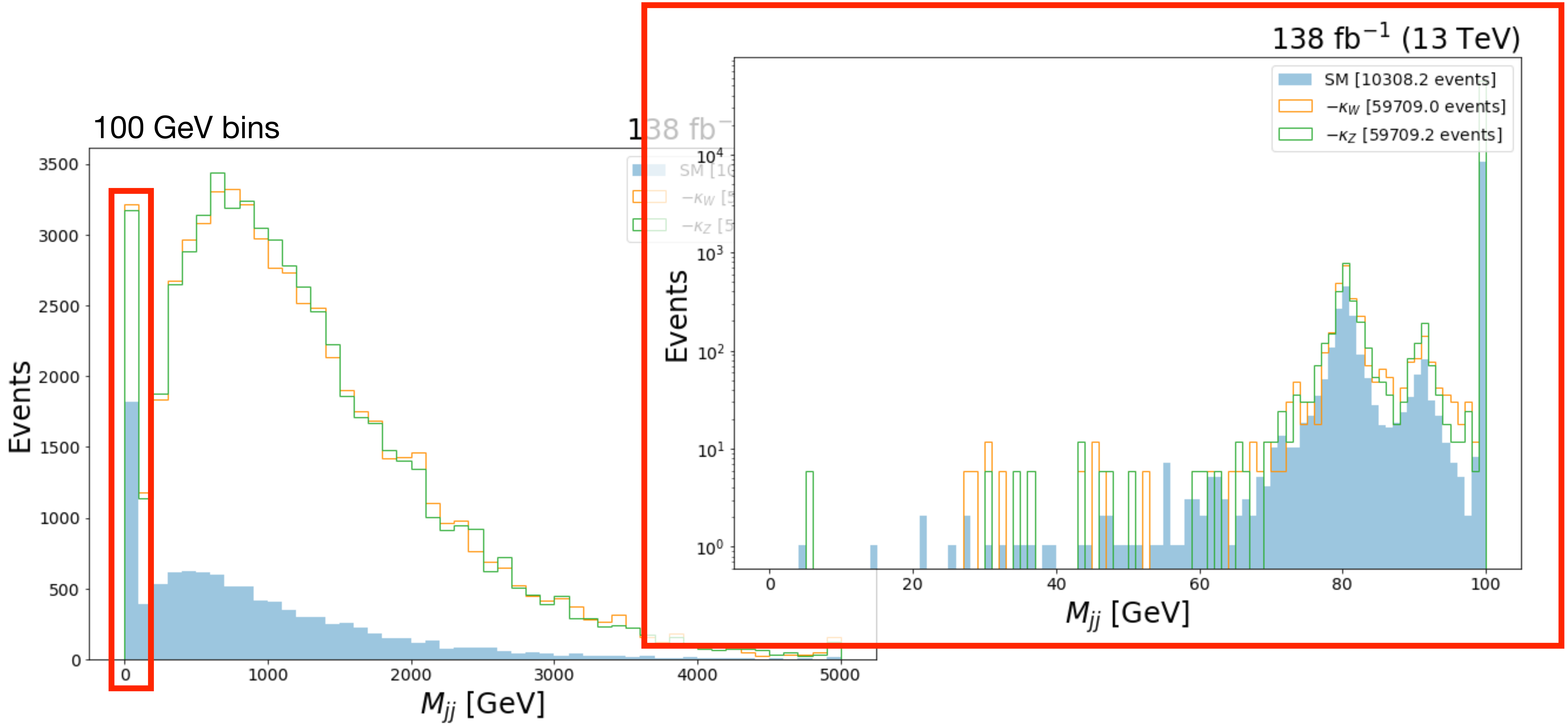
LHE VBS Kinematics





VBS-like signal: large $|\Delta\eta_{jj}|$, M_{jj}

LHE VBS Kinematics



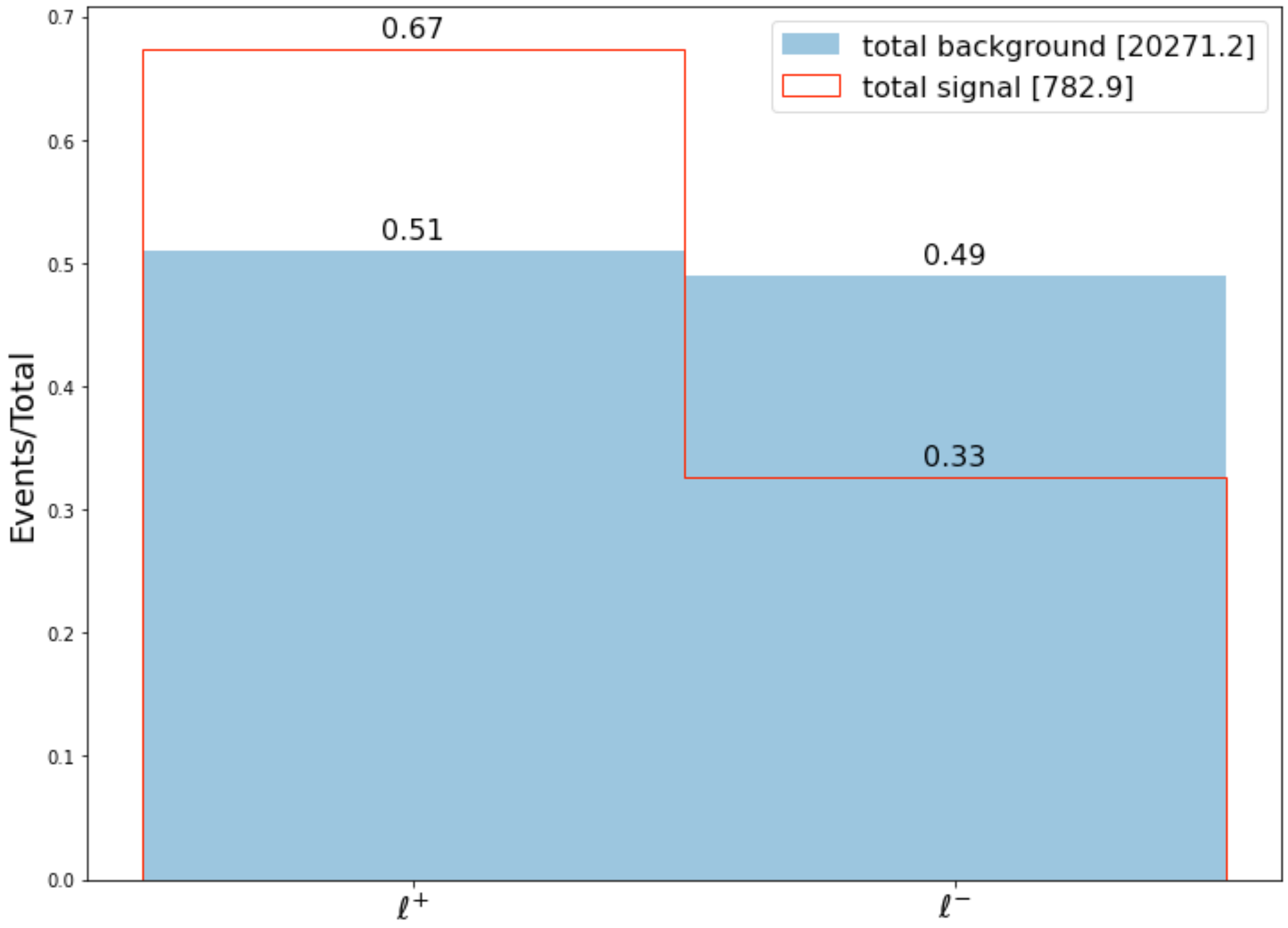
Visible W and Z peaks from pp → WHV (V→jj)

VBS WH Cutflow

	VHToNonbb		VV/VVV/VBSWZ		W+Jets		TTbar+X		TTbar1L		TTbar2L		TotalBkg		VBSWH (kW < 0)	
cut	raw	wgt	raw	wgt	raw	wgt	raw	wgt	raw	wgt	raw	wgt	raw	wgt	raw	wgt
Bookkeeping	1637	167.20	296595	1682.81	430128	32660.21	2960671	3675.78	9439330	424281.50	962161	34177.24	14090522	496644.74	399989	59598.09
SelectLeptons	1637	167.20	296595	1682.81	430128	32660.21	2960671	3675.78	9439330	424281.50	962161	34177.24	14090522	496644.74	399989	59598.09
Has1TightLep	1637	167.20	296595	1682.81	430128	32660.21	2960671	3675.78	9439330	424281.50	962161	34177.24	14090522	496644.74	73493	10993.58
LepPtGt40	1330	134.51	236196	1328.09	331677	26056.05	2229901	2744.47	7315482	328311.10	717050	25411.25	10831636	383985.47	60726	9071.35
SelectFatJets	1330	134.51	236196	1328.09	331677	26056.05	2229901	2744.47	7315482	328311.10	717050	25411.25	10831636	383985.47	60726	9071.35
Geq1FatJet	1317	134.75	232351	1310.98	324702	25501.02	2195215	2707.30	7268198	326174.50	665489	23568.50	10687272	379397.05	22021	3281.05
SelectHbbFatJet	1317	134.75	232351	1310.98	324702	25501.02	2195215	2707.30	7268198	326174.50	665489	23568.50	10687272	379397.05	22021	3281.05
PNetHbbScoreGt0p9	524	52.86	99990	692.78	81022	5851.27	849107	1005.49	2544585	114319.10	270367	9549.23	3845595	131470.73	8993	1344.32
SelectJetsNoHbbOverlap	524	52.86	99990	692.78	81022	5851.27	849107	1005.49	2544585	114319.10	270367	9549.23	3845595	131470.73	8993	1344.32
SelectVBSJetsMaxE	232	16.38	68999	220.91	61298	3224.51	759911	898.96	1926126	86676.40	188949	6690.97	3005515	97728.13	6210	929.32
MjjGt500detajjGt3	30	3.85	10365	35.87	11697	580.29	142157	169.63	399477	17930.50	43934	1551.03	607660	20271.17	5213	782.95

UC San Diego

Lepton Charge Asymmetry



CMS

UC San Diego

VBS WH BDT: XGBoost Configuration

Param	Value	Description
`objective`	binary:logistic	the learning task and the corresponding learning objective binary:logistic = 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
`base_score`	0.5	The initial prediction score of all instances, global bias
`scale_pos_weight`	6.76	Control the balance of positive and negative weights, useful for unbalanced classes; a typical value to consider: sum(negative instances) / sum(positive instances)

CMS

UC San Diego