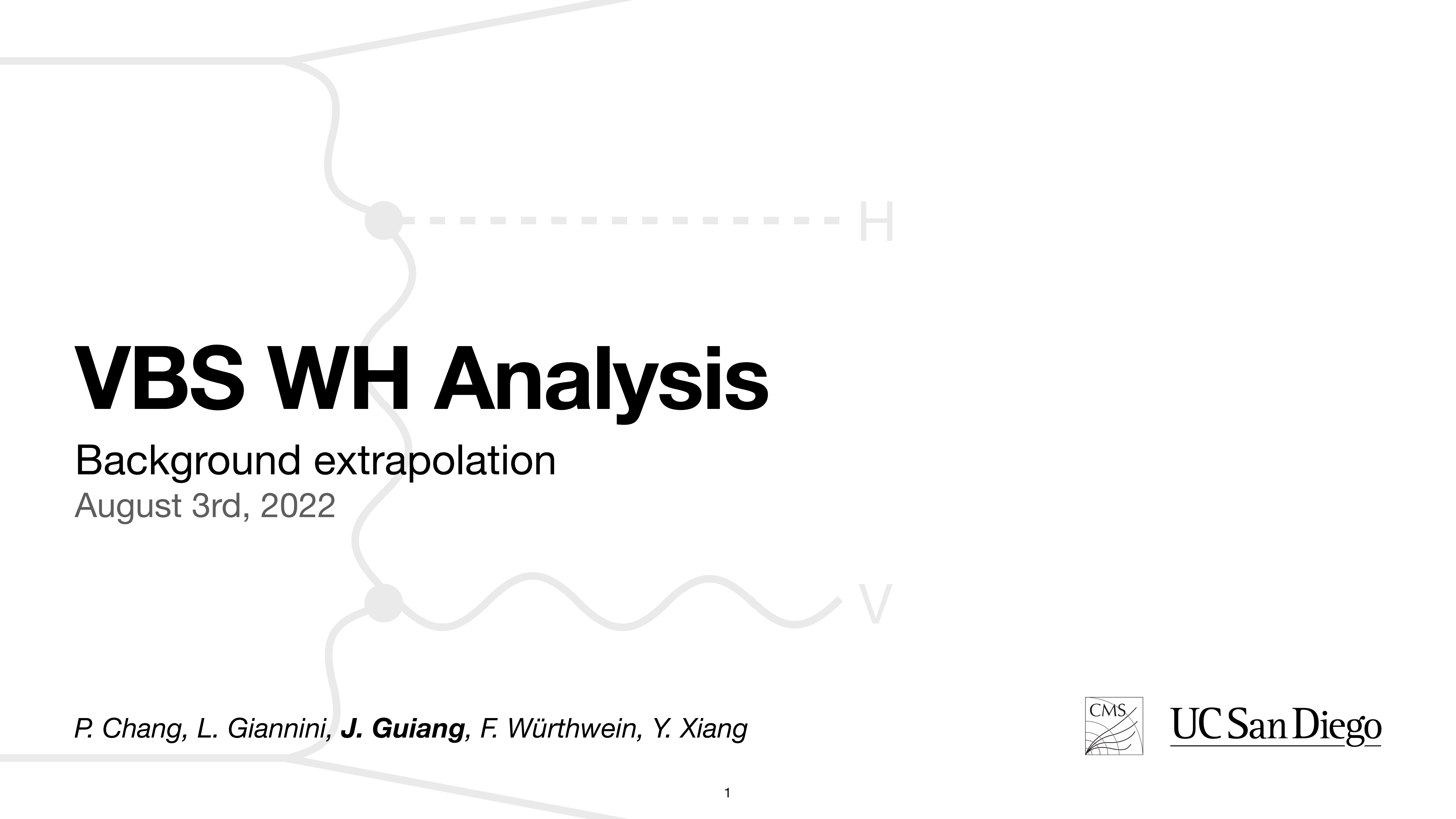


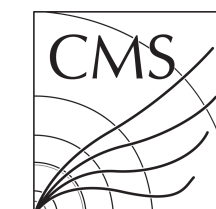
A decorative Feynman diagram in the background. It features a vertical line on the left with two vertices marked by grey circles. From the top vertex, a dashed line extends horizontally to the right, ending with the letter 'H'. From the bottom vertex, a wavy line extends horizontally to the right, ending with the letter 'V'.

VBS WH Analysis

Background extrapolation

August 3rd, 2022

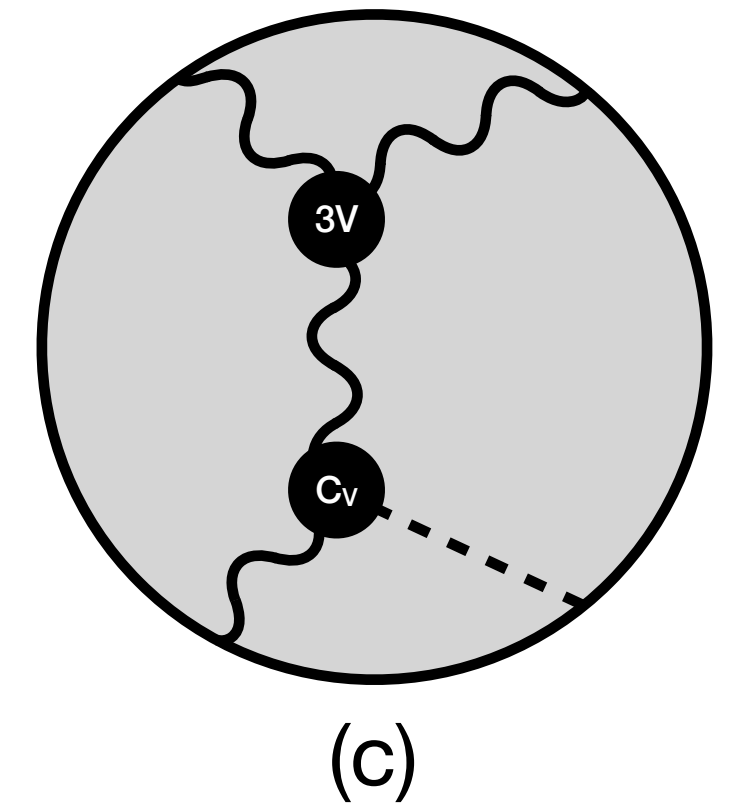
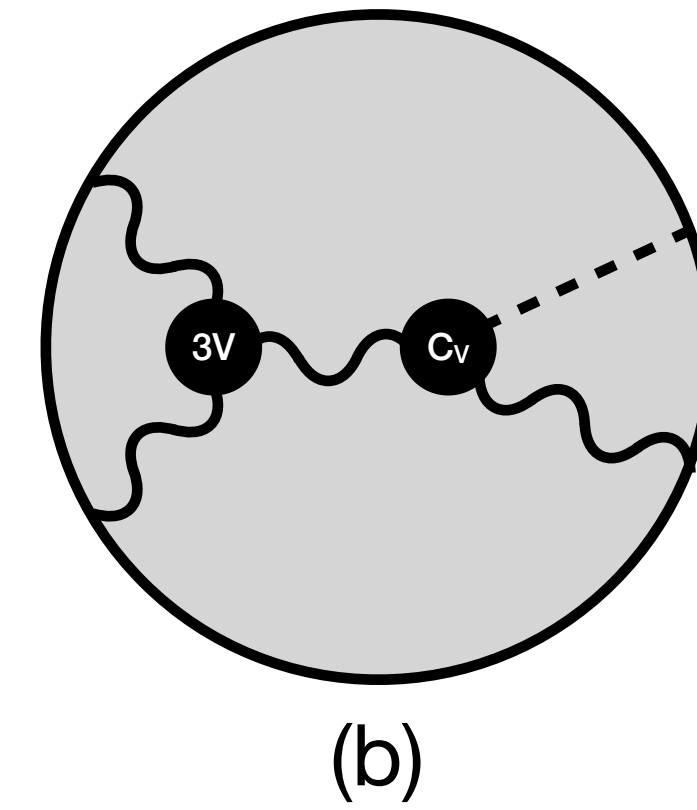
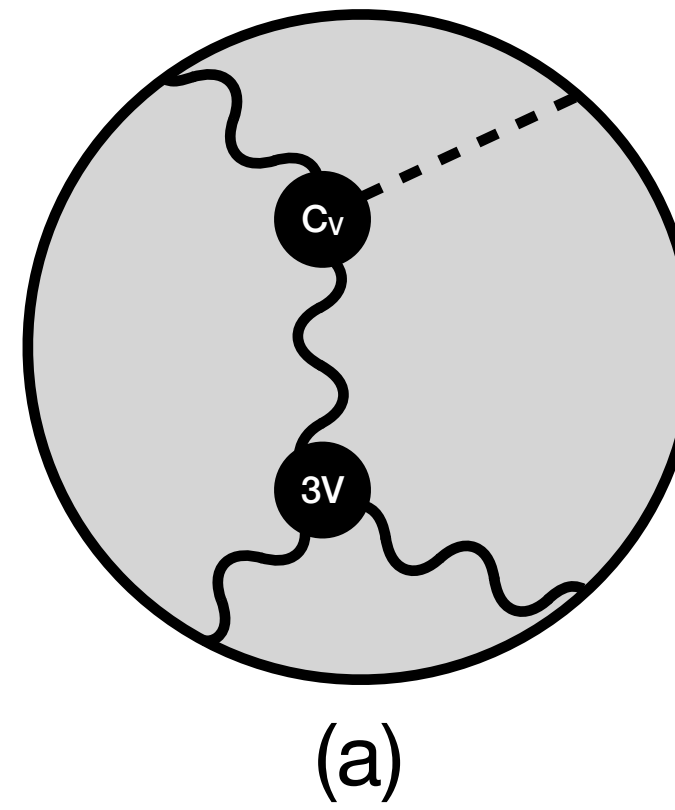
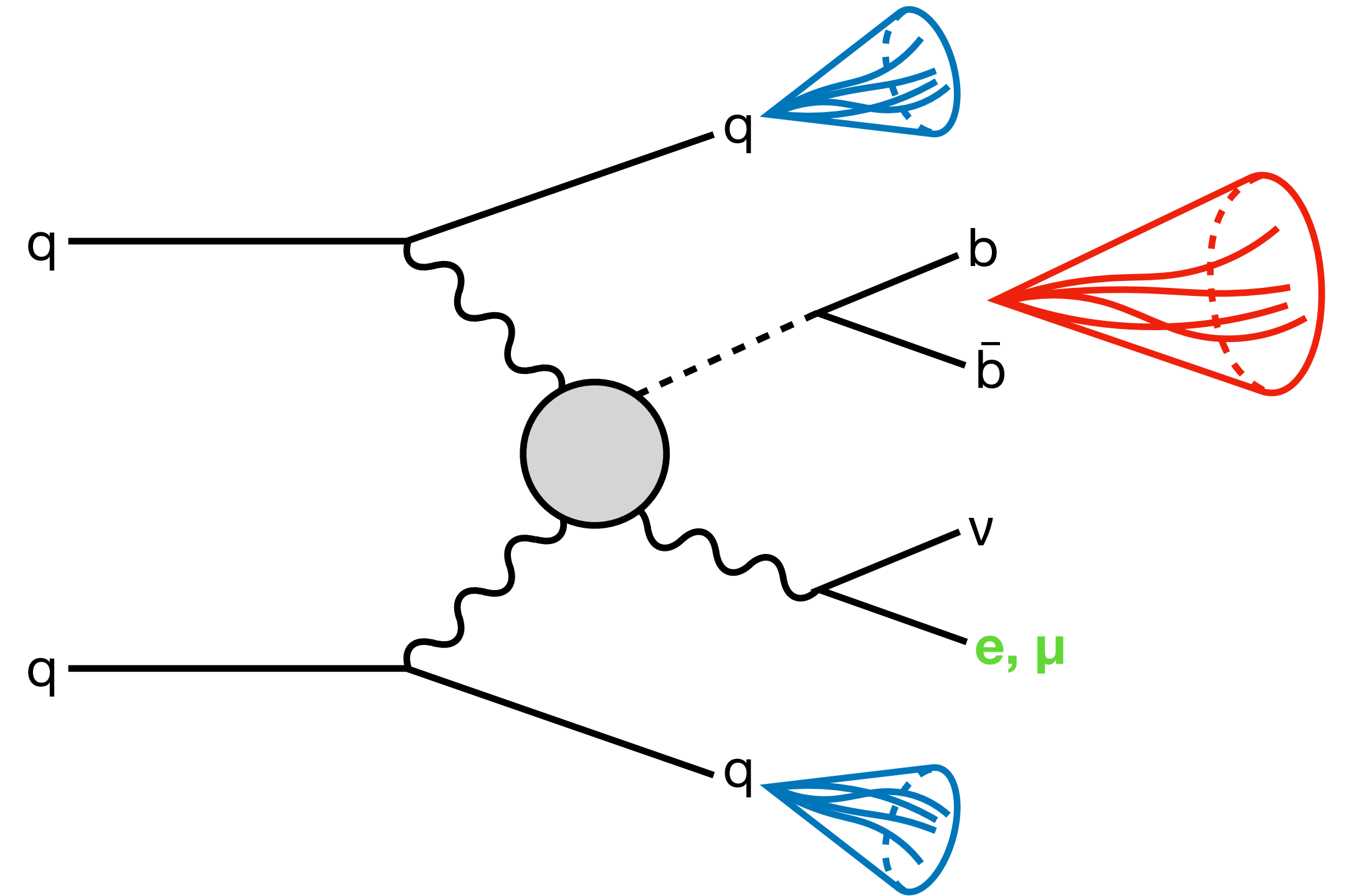
*P. Chang, L. Giannini, **J. Guiang**, F. Würthwein, Y. Xiang*



UC San Diego

Target Final State

- Targeting **VBS** $WH \rightarrow \ell \nu b \bar{b}$
- Sensitive to $\kappa_V \rightarrow \lambda_{WZ} = \kappa_Z/\kappa_W$
 - Handle for ruling out $\lambda_{WZ} = -1$ (BSM)
- VBS WH BSM kinematics:
 - High- p_T H and W (high S_T)
 - VBS jets with large $\Delta\eta_{jj}$, M_{jj}



Overview

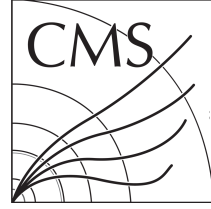
Yields scaled to $\text{lumi} \times \sigma$, rounded for readability

Cut		VH	VV/VVV/VBSWZ	W+Jets	SingleTop	TTbar+X	TTbar1L	TTbar2L	TotalBkg	Eff.*	VBSWH ($\kappa_W = -1$)	Eff.*
Presel.	Skim AND Postskim	232	6.5K	122K	7.6K	1K	86K	10K	233K	—	2K	—
	Object Selection	50	2K	47K	4.4K	654	50K	6.3K	111K	52%	1.1K	43%
	$M_{jj} > 500 \text{ GeV AND } \Delta\eta_{jj} > 3$	7	376	7.7K	994	124	10K	1.4K	21K	81%	920	17%
	$\text{max(PNet Xbb score)} > 0.3$	3	51	587	537	77	5.8K	382	7.4K	64%	608	34%
	No b-tagged (med.) AK4 jets	2	43	484	194	20	1.7K	110	2.5K	88%	579	37%
SR	$M_{jj} > 1.5 \text{ TeV AND } \Delta\eta_{jj} > 4$	1	5	68	31	3	297	22	427	83%	322	44%
	$S_T > 1.2 \text{ TeV}$	0	1	11	6	0	25	2	45	89%	139	57%
	PNet Xbb vs. QCD > 0.9	0	0	1	0	0	2	1	4	91%	109	22%
	AND $M_{SD} < 150$											

- Reminder: Postskim includes $S_T > 800 \text{ GeV}$ (full details in backup)
- Explored an ABCD extrapolation strategy that was viable in MC, but had bad data/MC agreement
 - See next slide



*eff = 1 - after/before

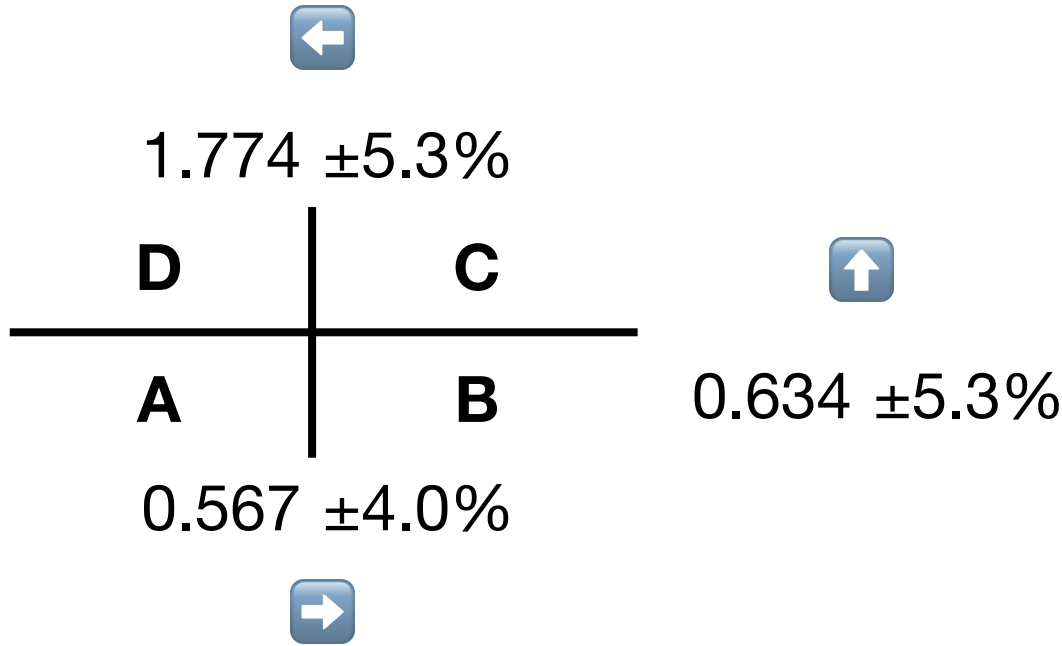


ABCD: $\Delta\eta_{jj}$ vs. M_{SD}

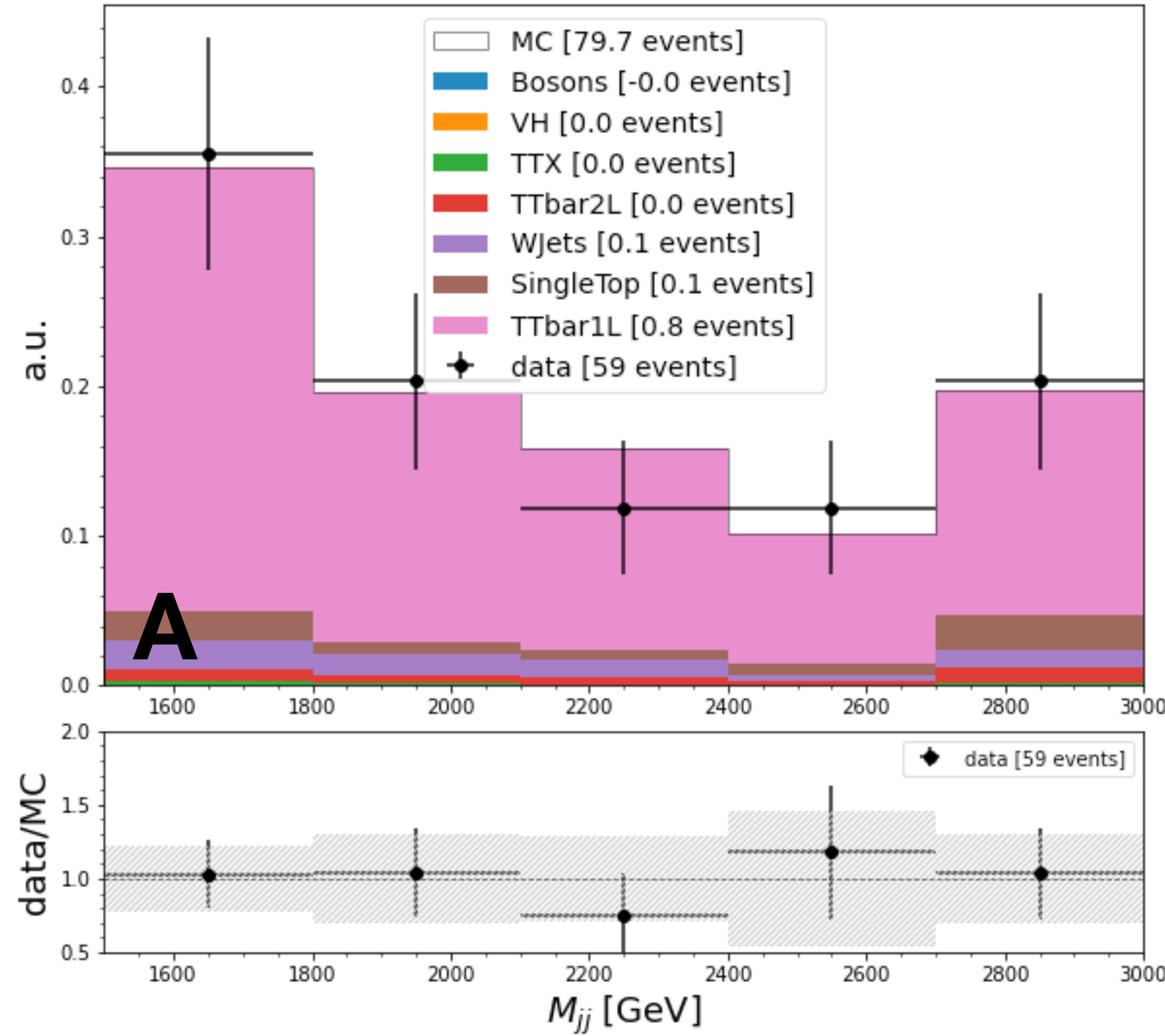
Presel. (w/out $\Delta\eta_{jj}$ cut) AND $M_{jj} > 1500$ GeV AND PNet Xbb > 0.9

Yields scaled to $lumix\sigma$, rounded for readability

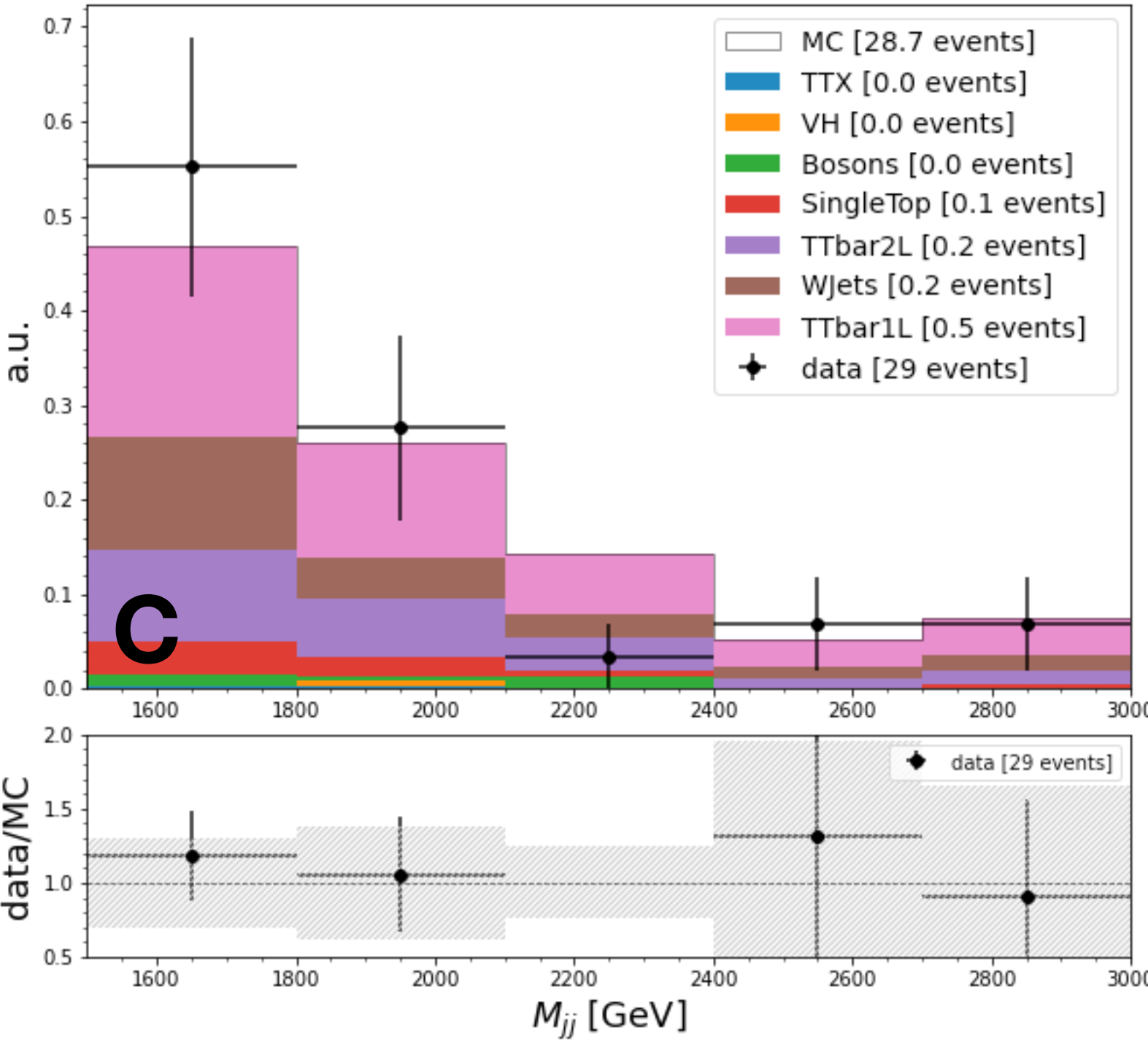
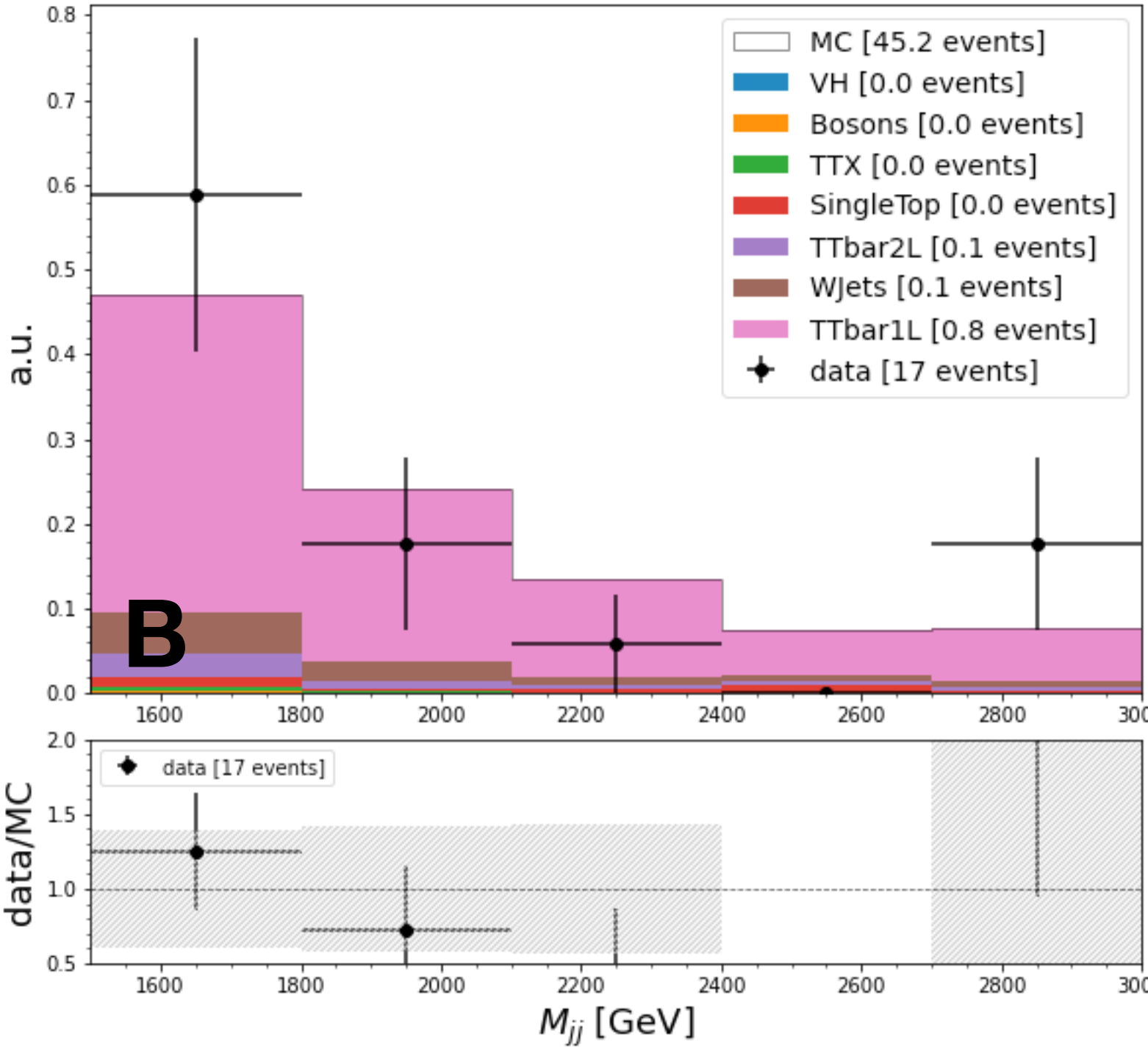
Cut	Region	Bkg. (raw)	Bkg. (wgt)	Bkg. Err.*	Sig. (raw)	Sig. (wgt)	Sig. Err.*
$ \Delta\eta_{jj} > 4$ AND $M_{SD} \geq 150$ GeV	A	2902	79.68	2.02	43	6.30	1.06
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} \geq 150$ GeV	B	2662	45.21	1.42	2	0.50	0.35
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} < 150$ GeV	C	2481	28.66	1.21	13	1.57	0.50
$ \Delta\eta_{jj} > 4$ AND $M_{SD} < 150$ GeV (~SR**)	D	2492	50.84	1.65	1658	244.98	6.70



59 data vs 80 MC



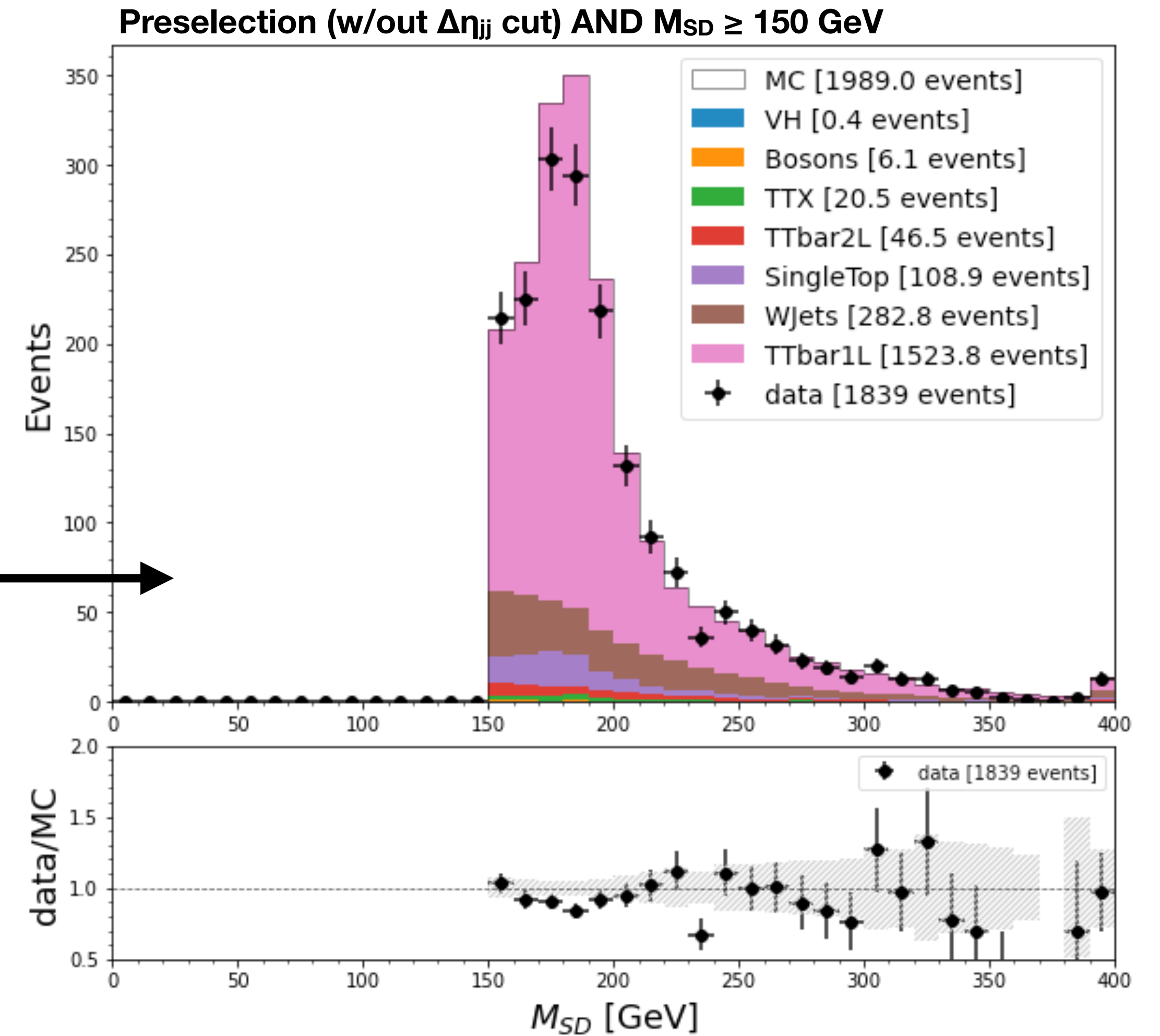
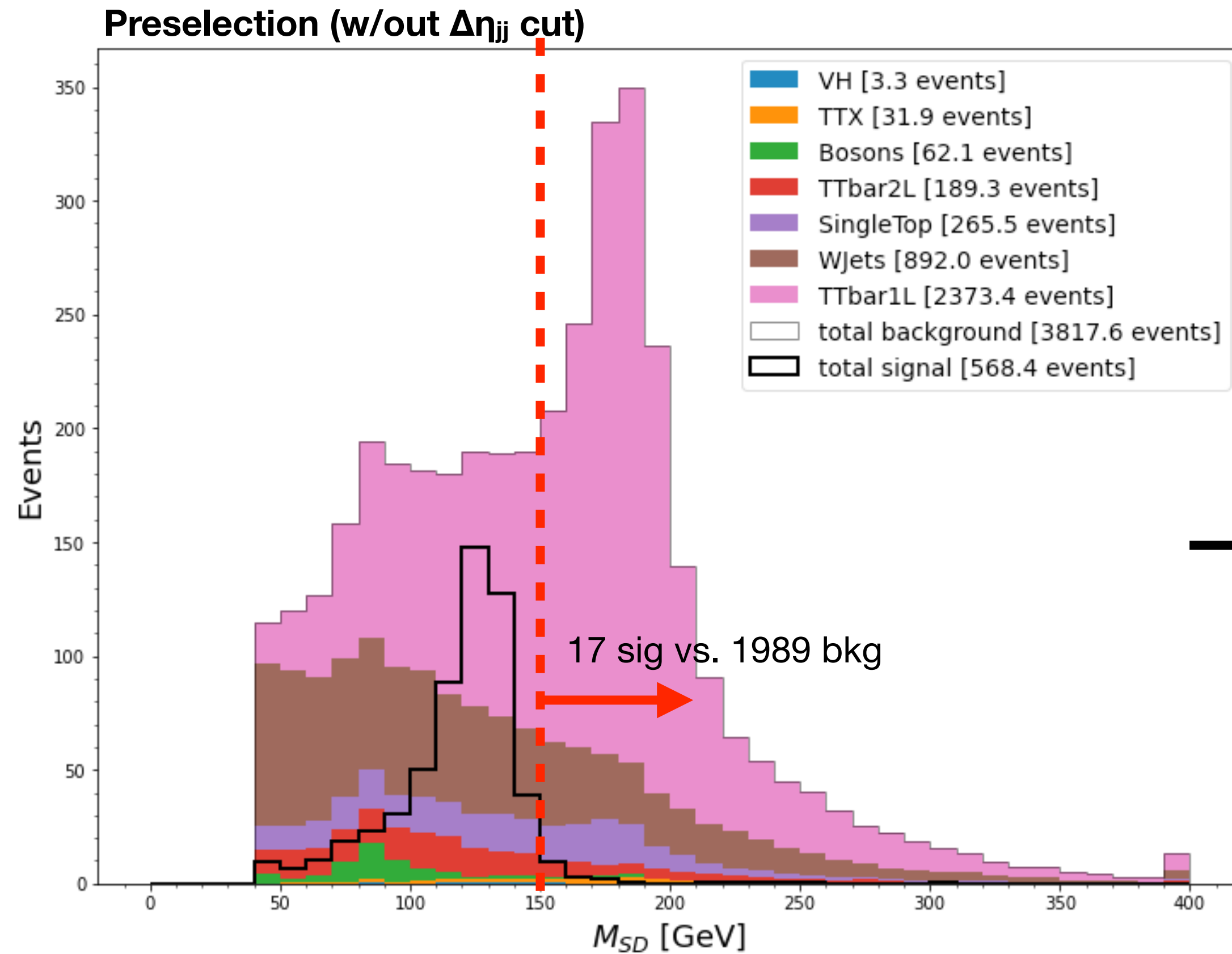
17 data vs 45 MC



*err = $\sqrt{(\sum_i w_i^2)}$

**Specifically, this is the SR with a looser S_T cut ($S_T > 800$ GeV)

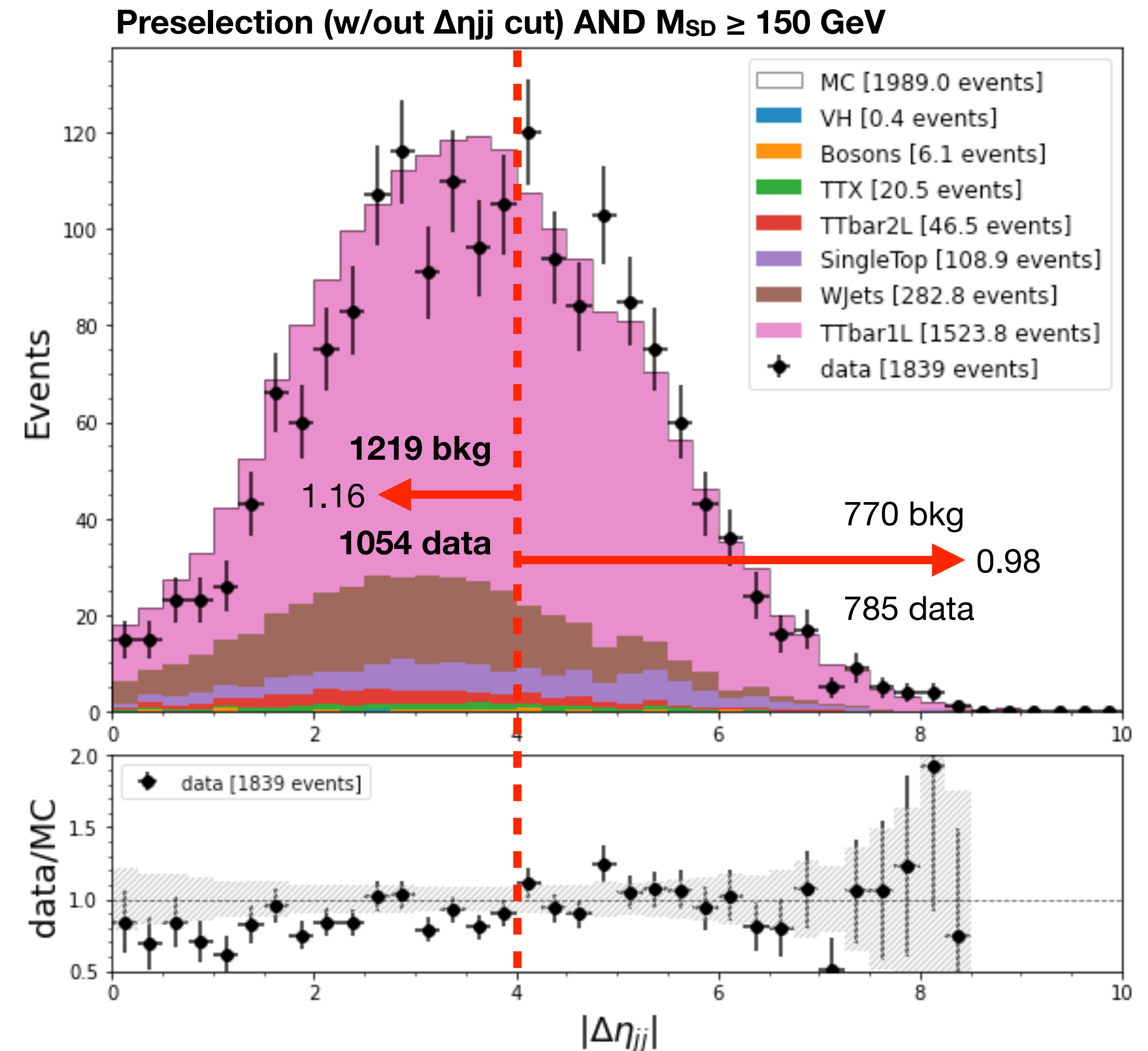
Walking Towards the Discrepancy



Agreement looks good so far...

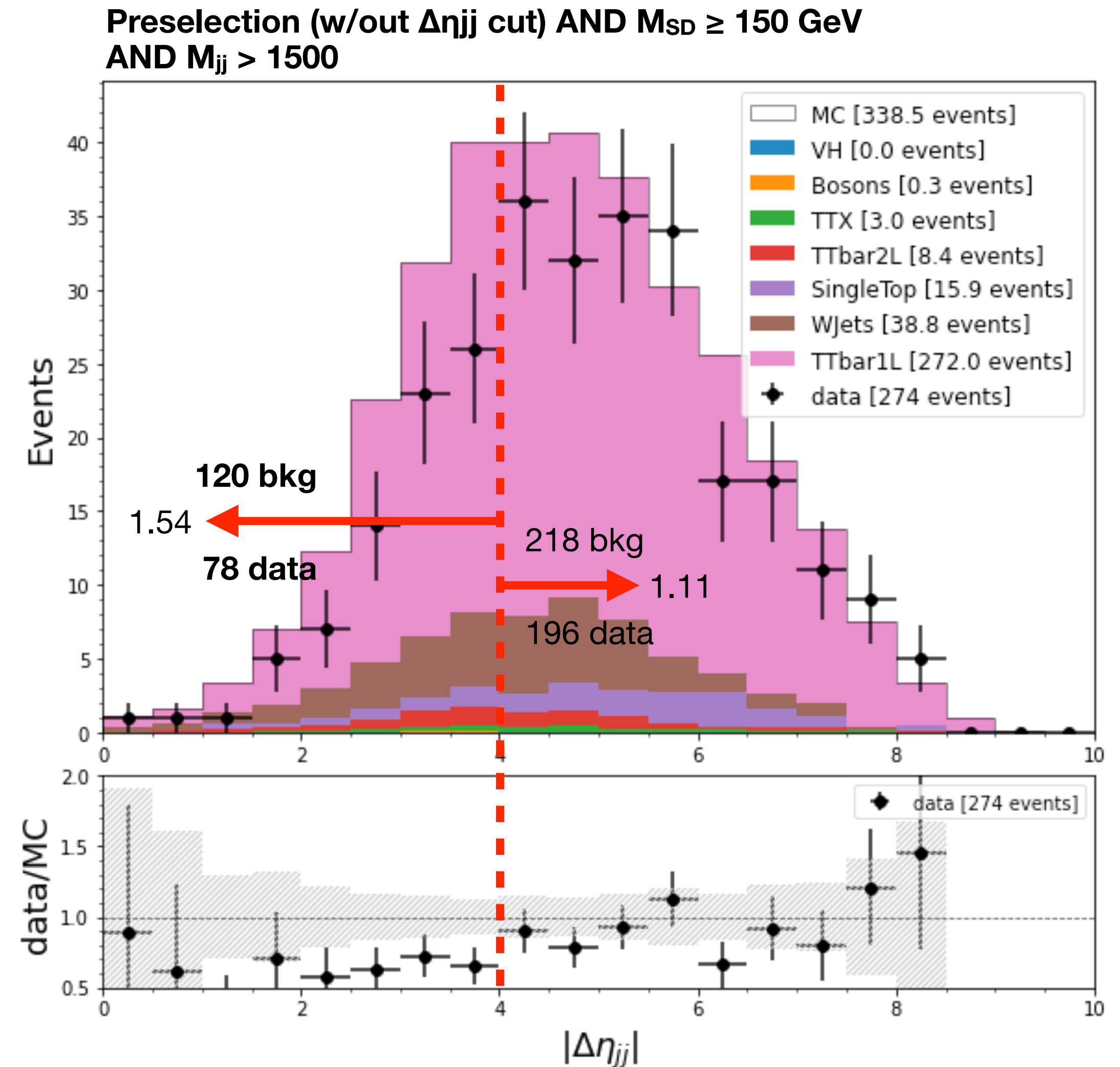
Walking Towards the Discrepancy

- Bkg. MC vs. data shown on plot on either side of $\Delta\eta_{jj}$ cut
- Only lepton scale factors are applied
 - Bkg. MC weighted to $\text{lumi} \times \sigma$
- **Significant excess of MC for $|\Delta\eta_{jj}| \leq 4$**
- **Data $\sim$ MC for $|\Delta\eta_{jj}| > 4$**
- **Already looks somewhat like Region A and Region B in ABCD**



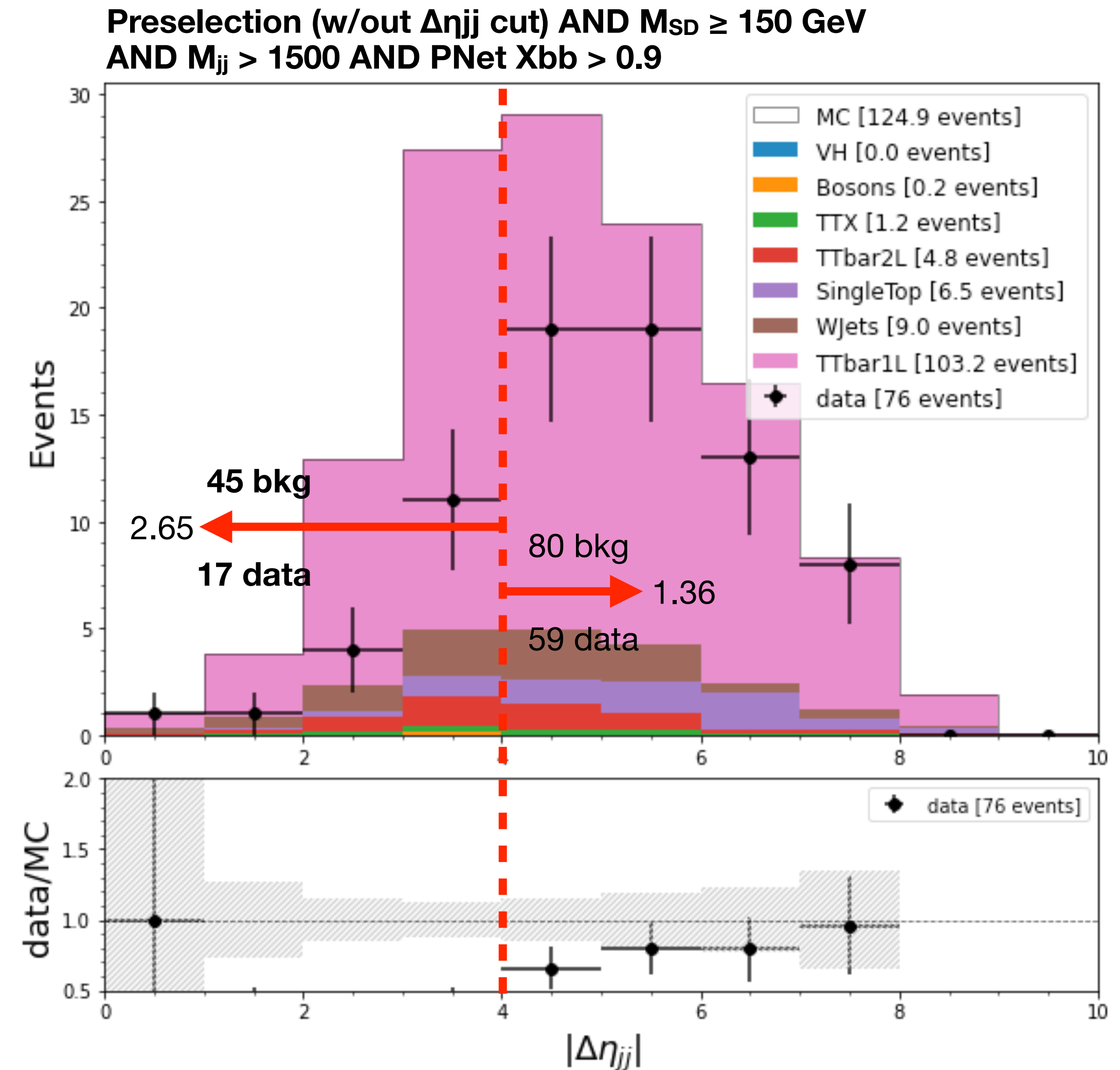
Walking Towards the Discrepancy

- Bkg. MC vs. data shown on plot on either side of $\Delta\eta_{jj}$ cut
- Only lepton scale factors are applied
 - Bkg. MC weighted to $\text{lumi} \times \sigma$
- **Significant excess of MC for $|\Delta\eta_{jj}| \leq 4$**
- **Now see excess of MC for $|\Delta\eta_{jj}| > 4$**
- **Now looks more like Region A and Region B in ABCD**

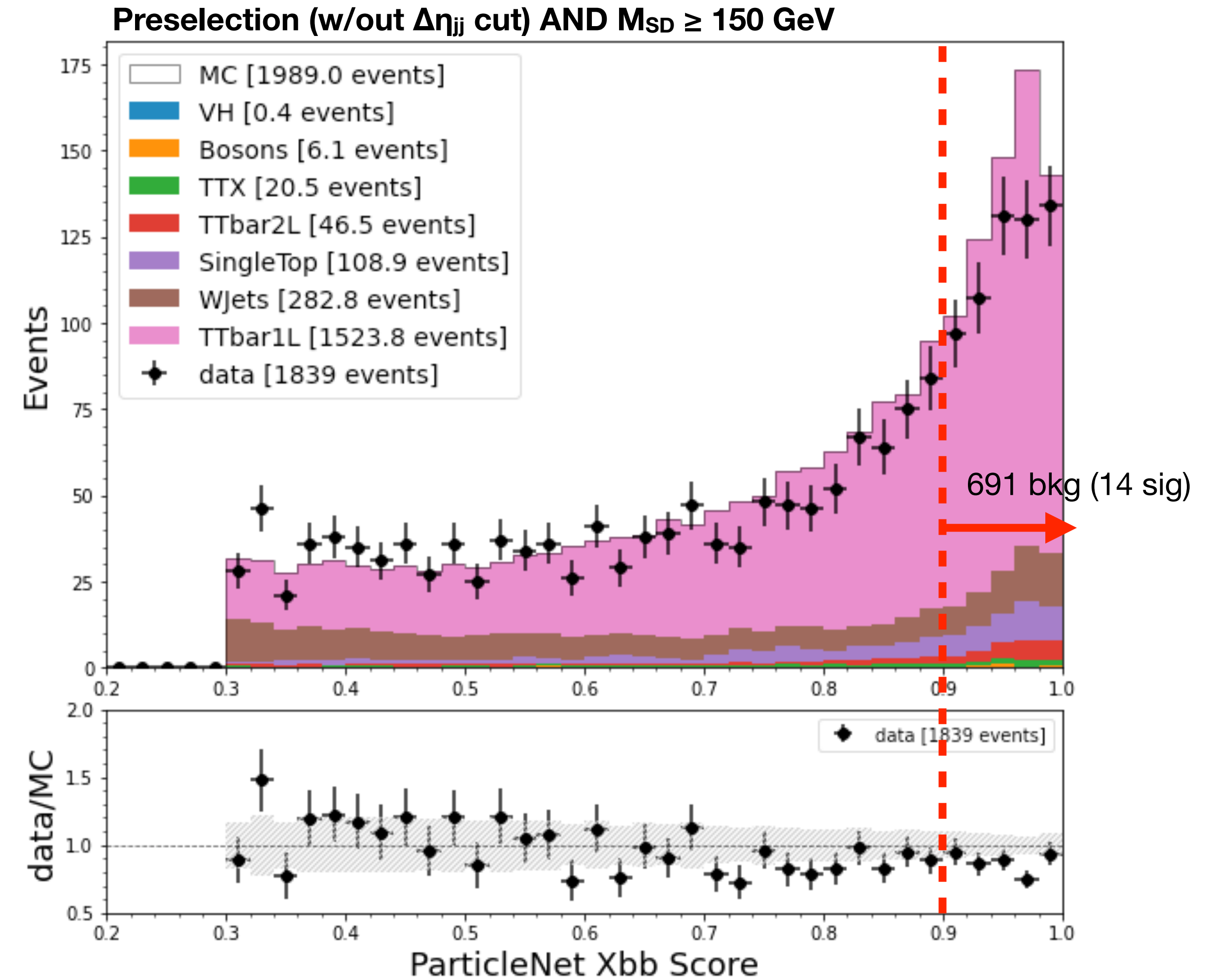
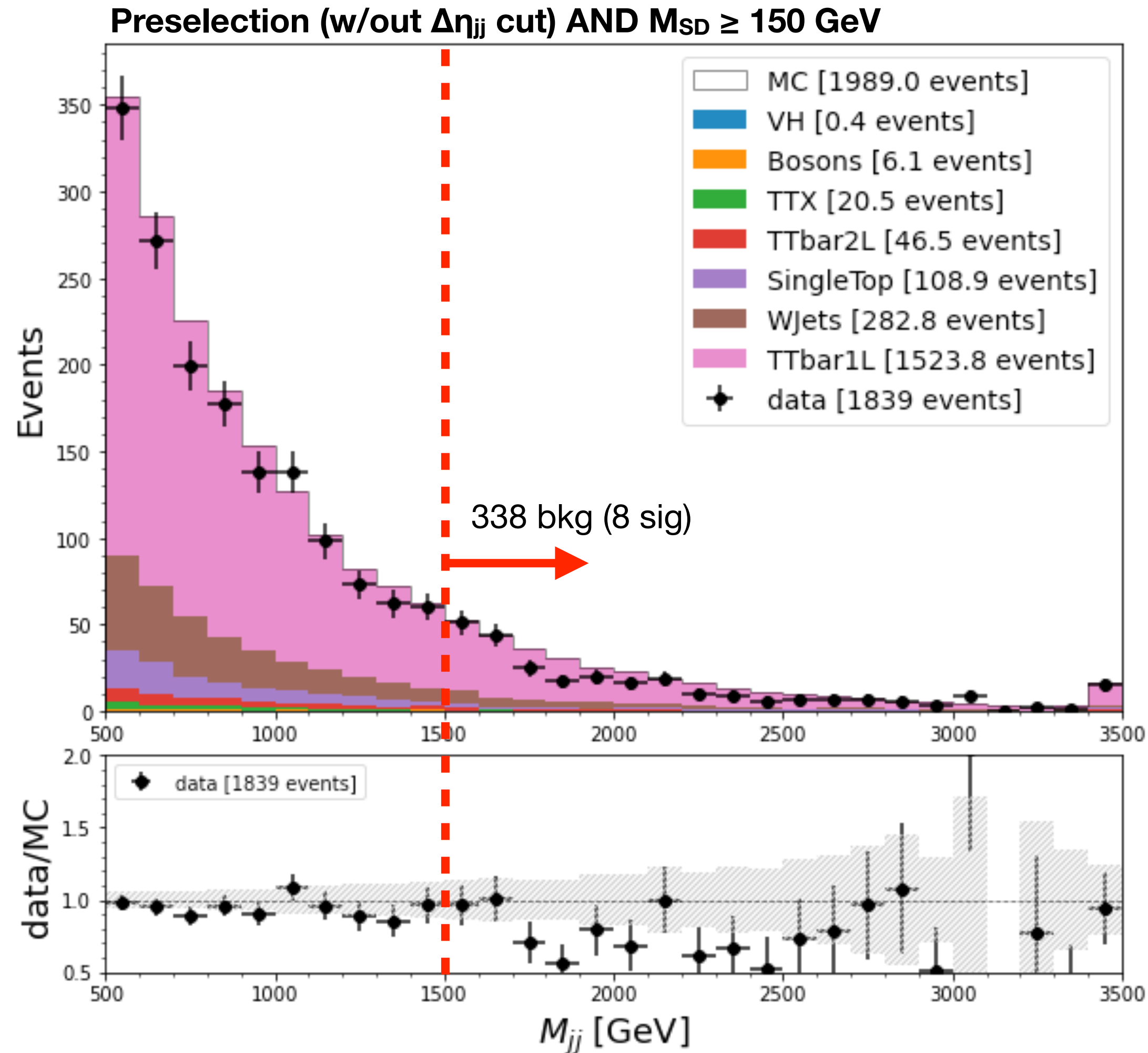


Walking Towards the Discrepancy

- Bkg. MC vs. data shown on plot on either side of $\Delta\eta_{jj}$ cut
- Only lepton scale factors are applied
 - Bkg. MC weighted to $\text{lumi} \times \sigma$
- **Significant excess of MC for $|\Delta\eta_{jj}| \leq 4$**
- **Now see excess of MC for $|\Delta\eta_{jj}| > 4$**
- **Arrived at Region A and Region B in ABCD**

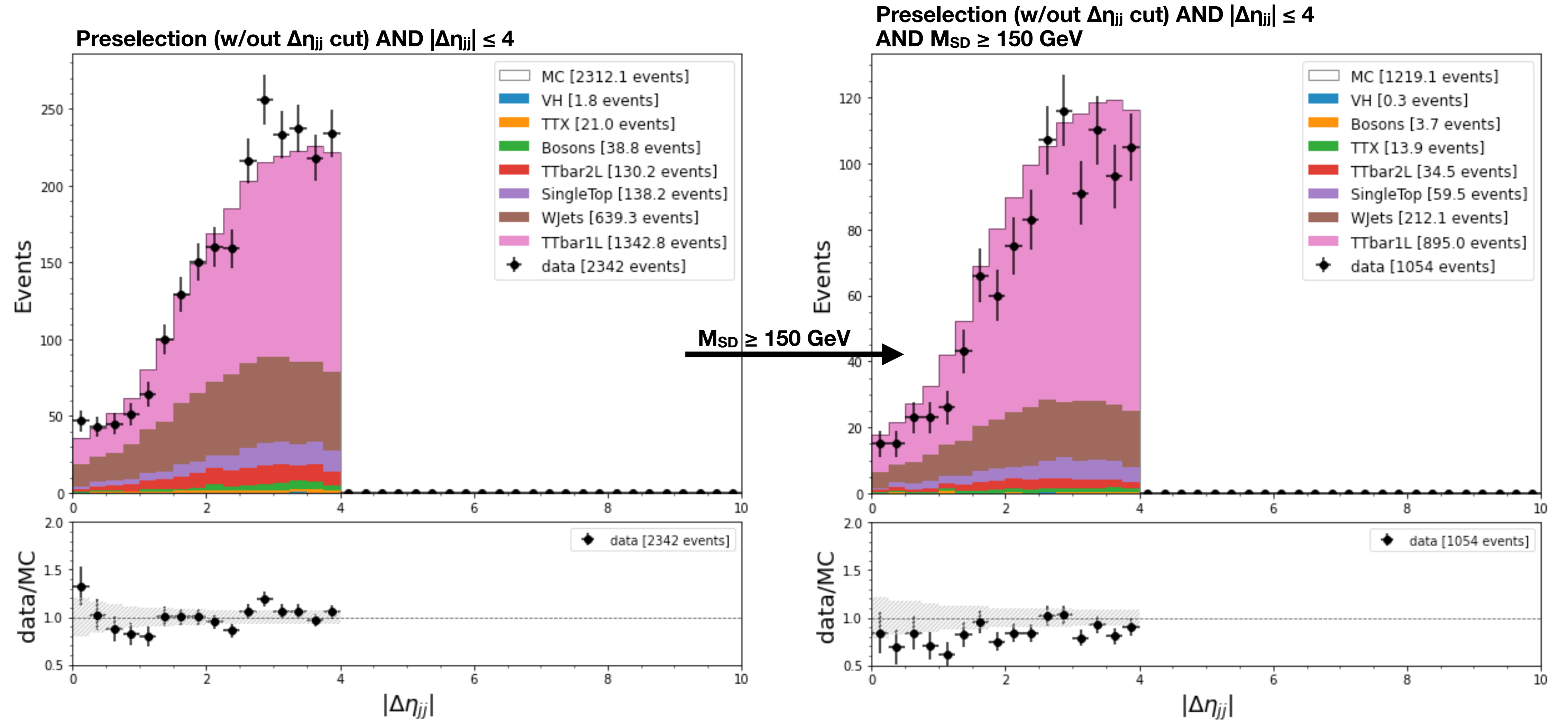


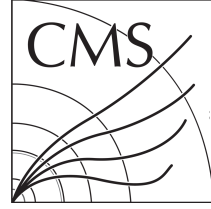
Walking Towards the Discrepancy



No clear data deficiency caused by either cut alone

Walking Towards the Discrepancy

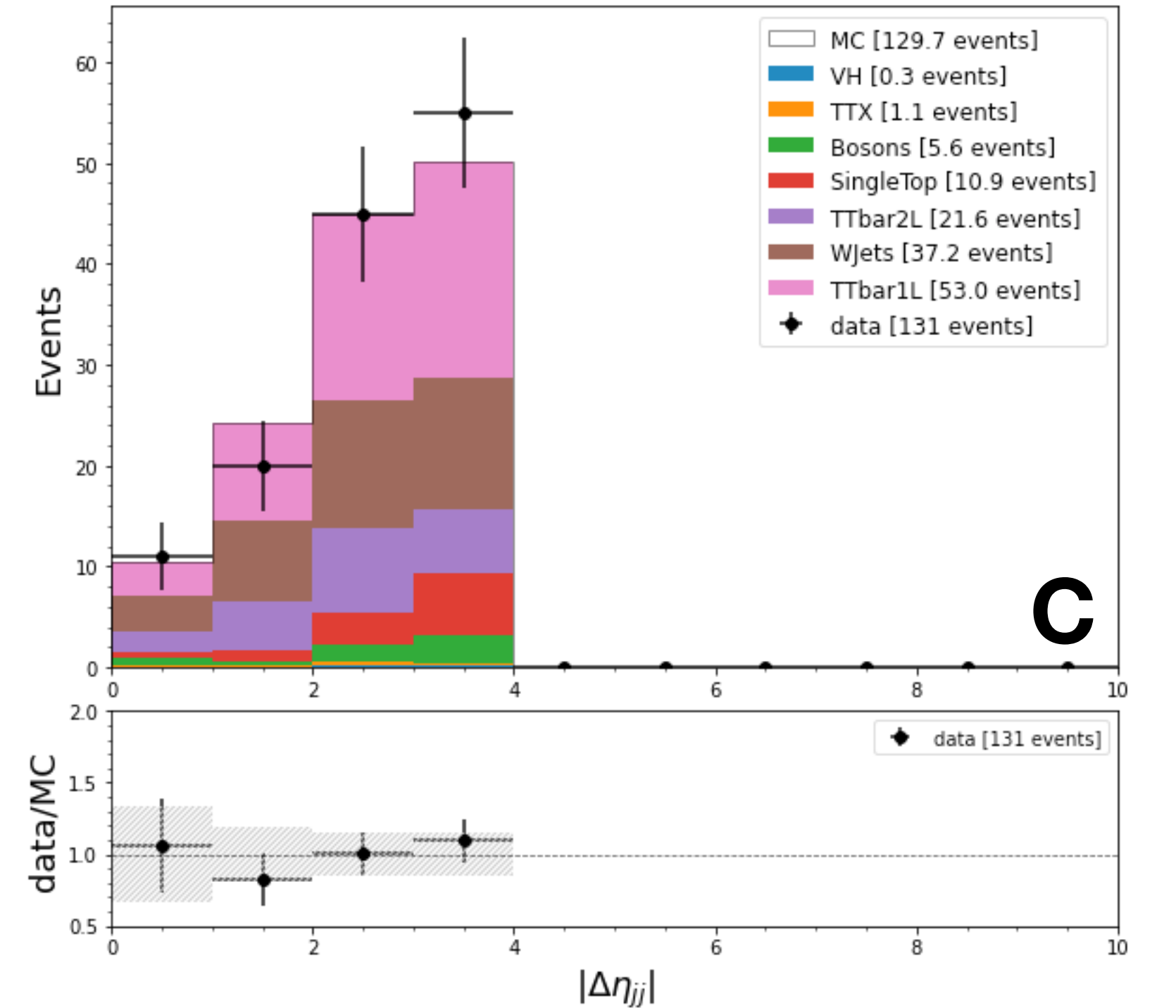
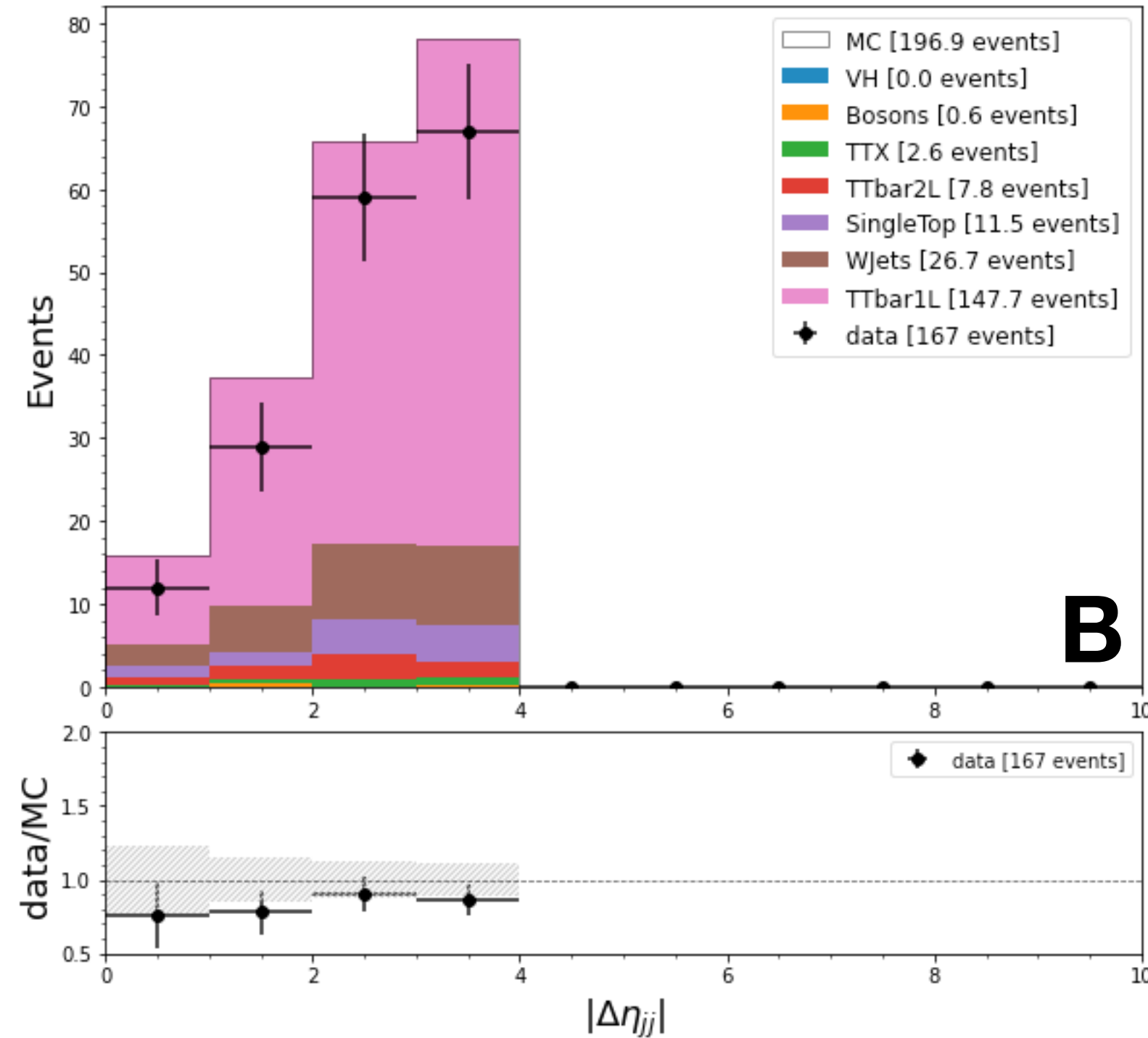
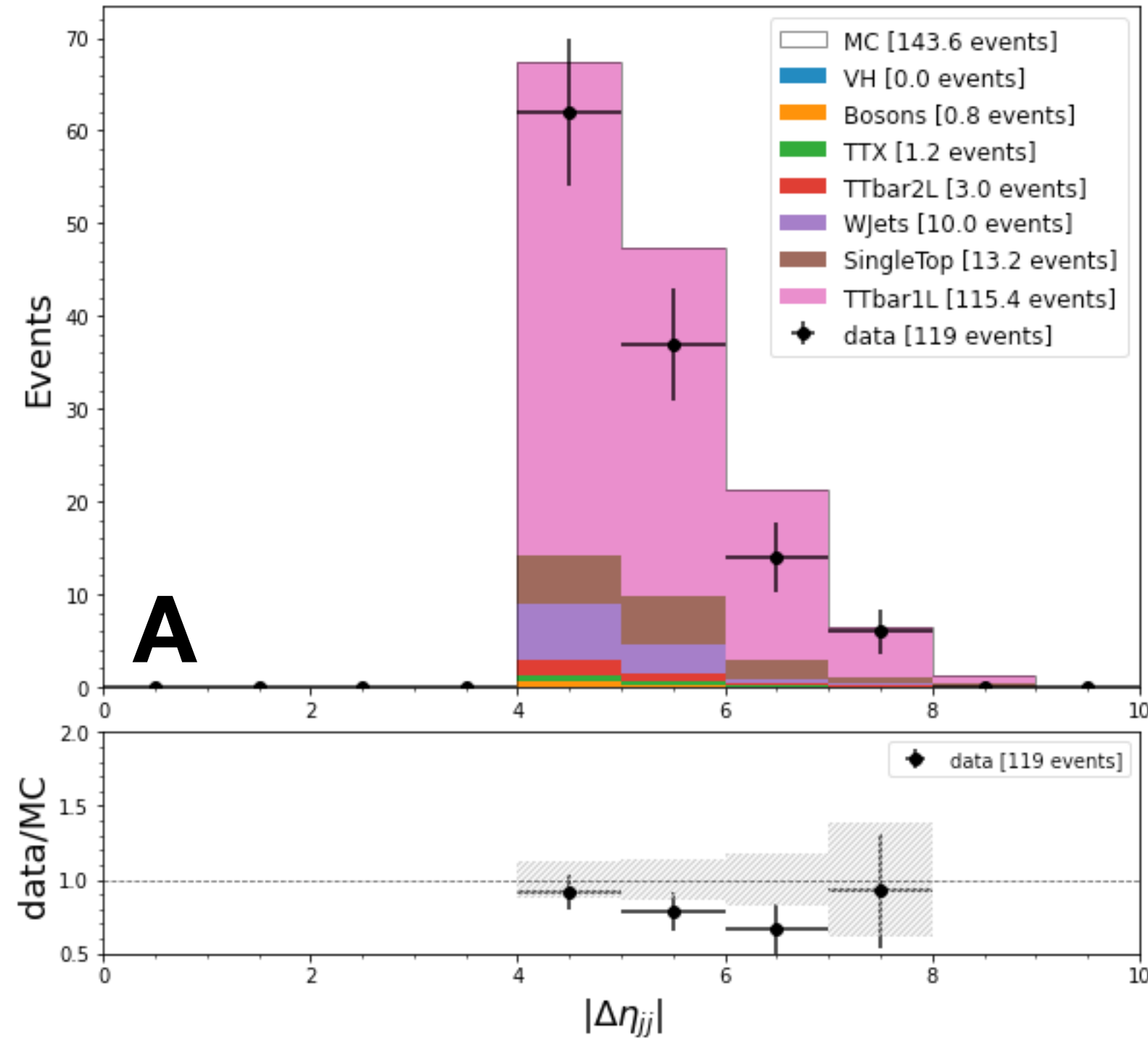
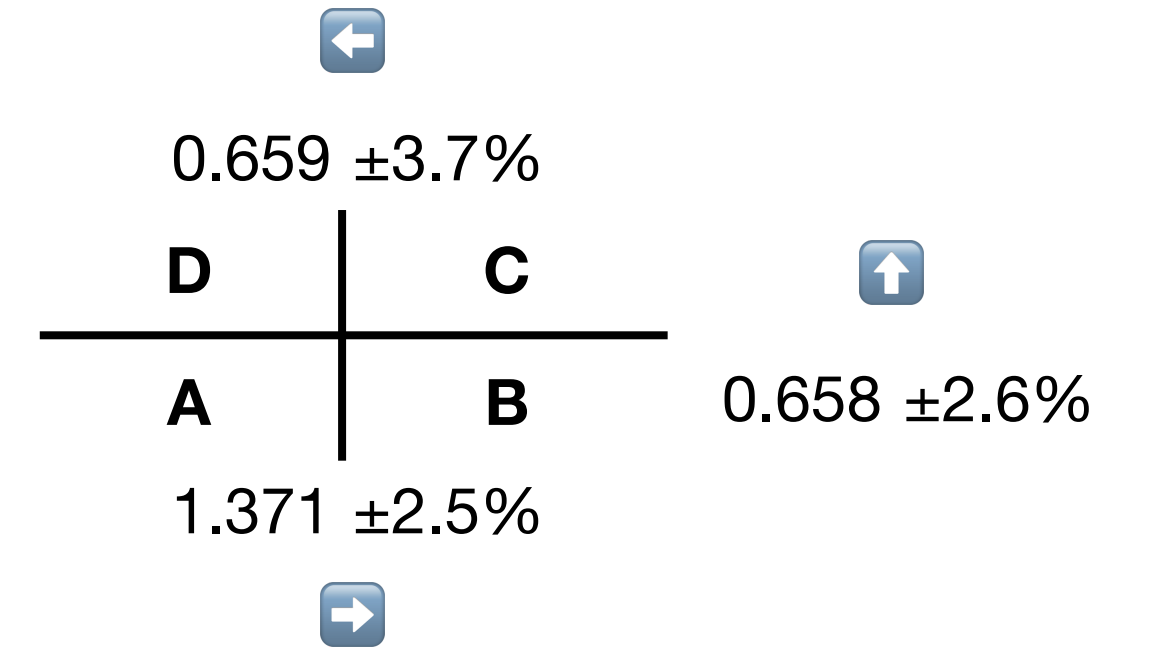




ABCD: $\Delta\eta_{jj}$ vs. M_{SD}

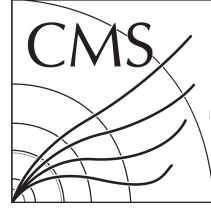
Presel. (w/out $\Delta\eta_{jj}$ cut) AND $M_{jj} > 600$ GeV AND $S_T > 900$ GeV AND $P_{Net} X_{bb} > 0.9$

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$ \Delta\eta_{jj} > 4$ AND $M_{SD} \geq 150$ GeV	A	143.62	2.88	8.08	1.16	119	10.91
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} \geq 150$ GeV	B	196.94	3.08	0.93	0.45	167	12.92
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} < 150$ GeV	C	129.68	2.67	12.10	1.44	131	11.45
$ \Delta\eta_{jj} > 4$ AND $M_{SD} < 150$ GeV (~SR**)	D	85.42	2.58	336.33	7.79	—	—



*err = $\sqrt{(\sum_i w_i^2)}$

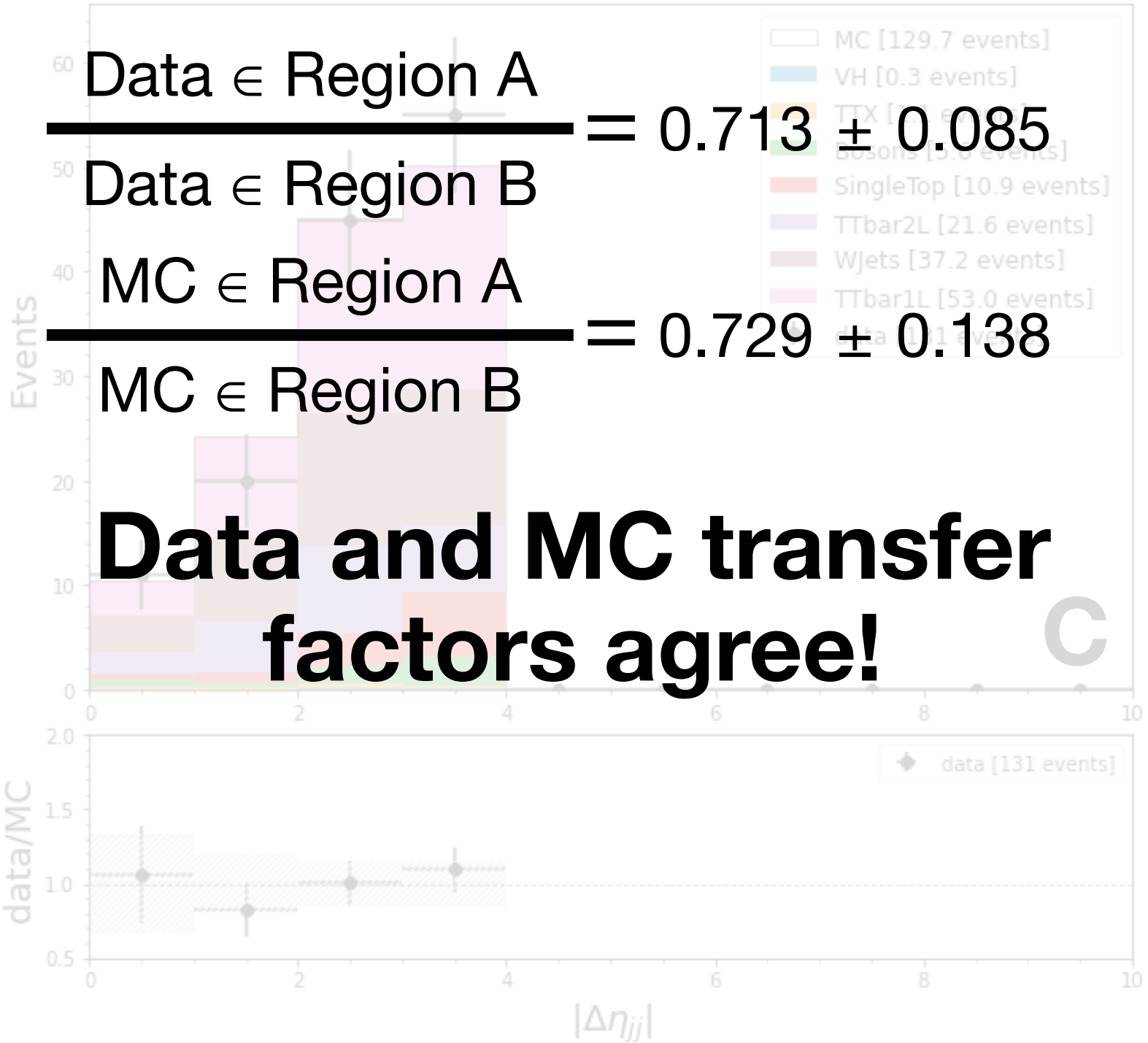
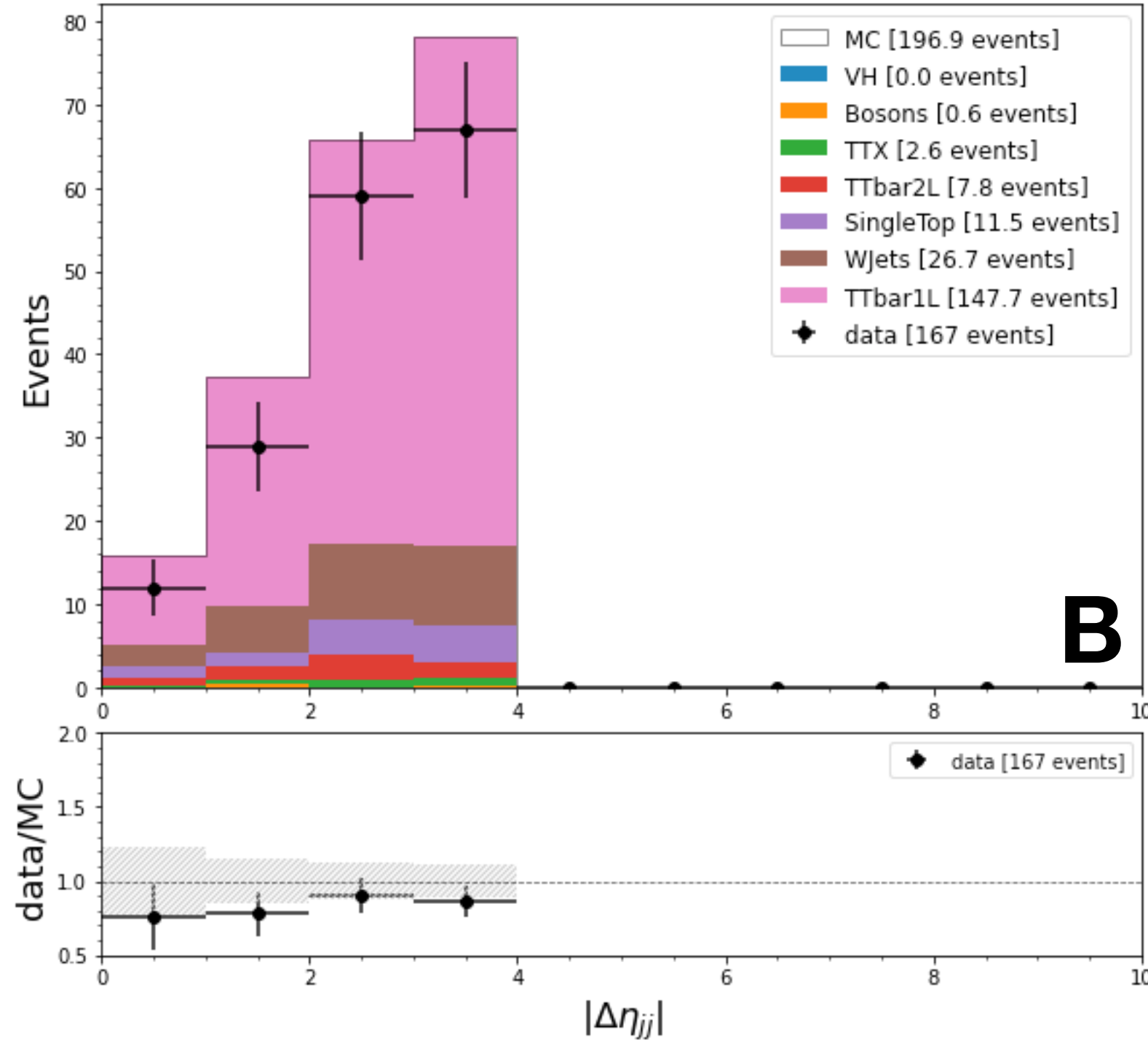
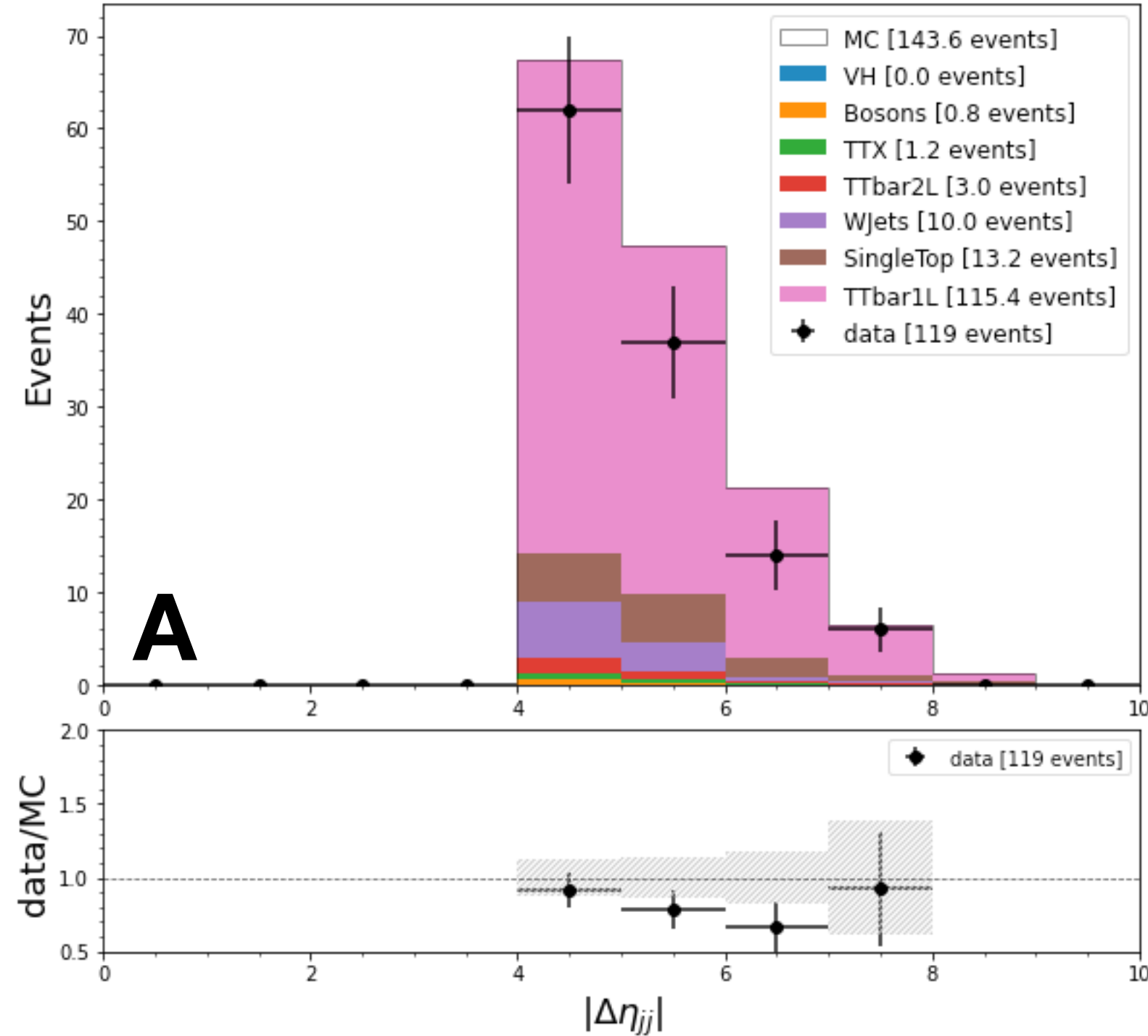
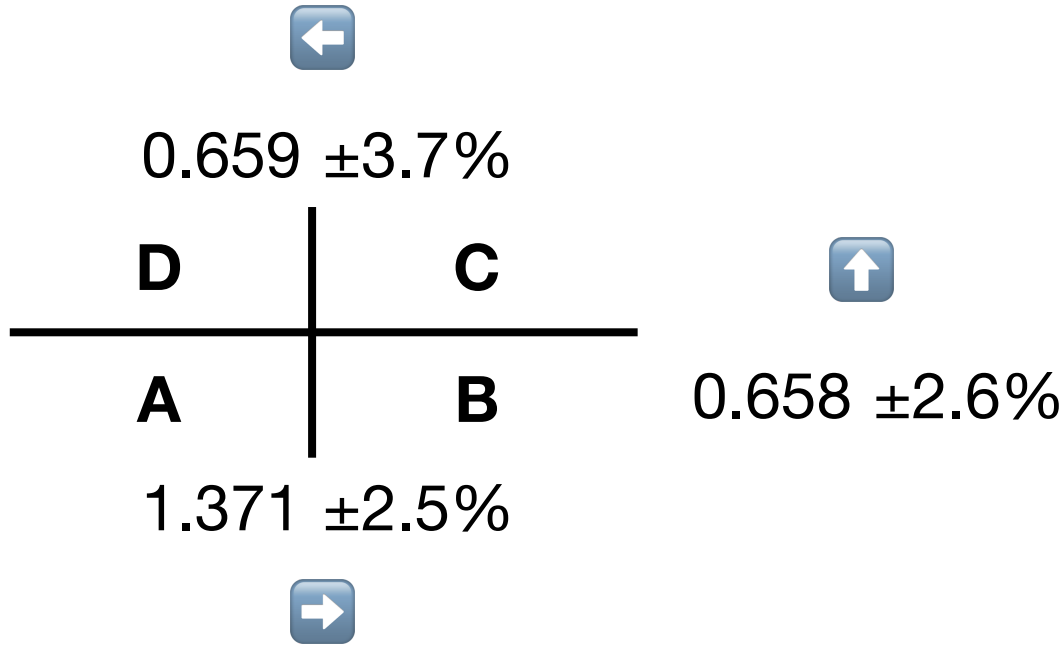
**Specifically, this is the SR with a looser S_T cut ($S_T > 800$ GeV)



ABCD: $\Delta\eta_{jj}$ vs. M_{SD}

Presel. (w/out $\Delta\eta_{jj}$ cut) AND $M_{jj} > 600$ GeV AND $S_T > 900$ GeV AND $P_{Net} X_{bb} > 0.9$

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$ \Delta\eta_{jj} > 4$ AND $M_{SD} \geq 150$ GeV	A	143.62	2.88	8.08	1.16	119	10.91
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} \geq 150$ GeV	B	196.94	3.08	0.93	0.45	167	12.92
$ \Delta\eta_{jj} \leq 4$ AND $M_{SD} < 150$ GeV	C	129.68	2.67	12.10	1.44	131	11.45
$ \Delta\eta_{jj} > 4$ AND $M_{SD} < 150$ GeV (~SR**)	D	85.42	2.58	336.33	7.79	—	—



Data $\in$ Region A
 Data $\in$ Region B
 MC $\in$ Region A
 MC $\in$ Region B

$= 0.713 \pm 0.085$
 $= 0.729 \pm 0.138$

Data and MC transfer factors agree!

*err = $\sqrt{(\sum_i w_i^2)}$ for MC, $\sqrt{(\text{count})}$ for data
 **Specifically, this is the original SR with a looser S_T and M_{jj} cut

Conclusions

- Poor data/MC agreement for region A of ABCD strategy seems to exist before M_{jj} and $\Delta\eta_{jj}$ cuts are applied
- Discrepancy in regions *A and B* is made worse by M_{jj} cut
 - Maybe M_{jj} vs. $\Delta\eta_{jj}$ correlation is not well-modeled?
 - M_{SD} and $\Delta\eta_{jj}$ are uncorrelated, but discrepancy starts after cutting on M_{SD}
- Cutting much looser on M_{jj} and compensating somewhat with S_T gives us a worse SR, but a viable background extrapolation
 - 336 sig. vs. 85 bkg.
 - If we take $\sqrt{336} \sim 18$ as 1σ , must fluctuate down by $> 10\sigma$ to reach 85 bkg.

Backup

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*

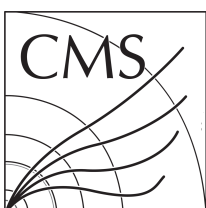
UC San Diego

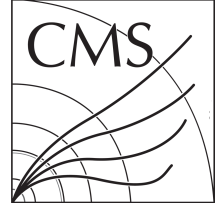
UNIVERSITY OF
Nebraska
Lincoln



UC SANTA BARBARA

UC San Diego





BSM Signal Models

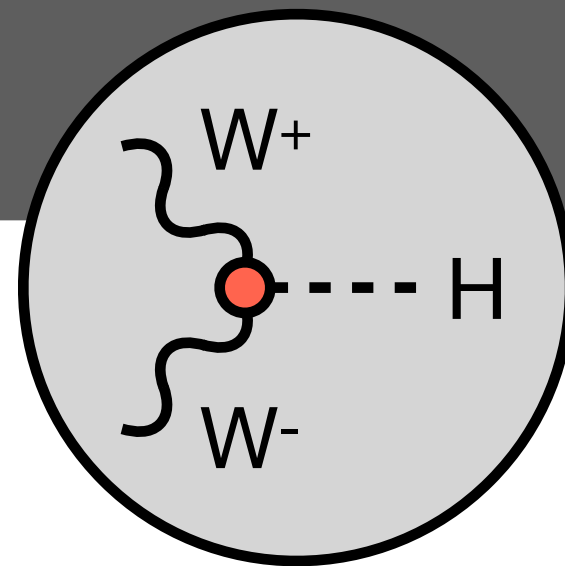
$$K_W = -1$$

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})
```



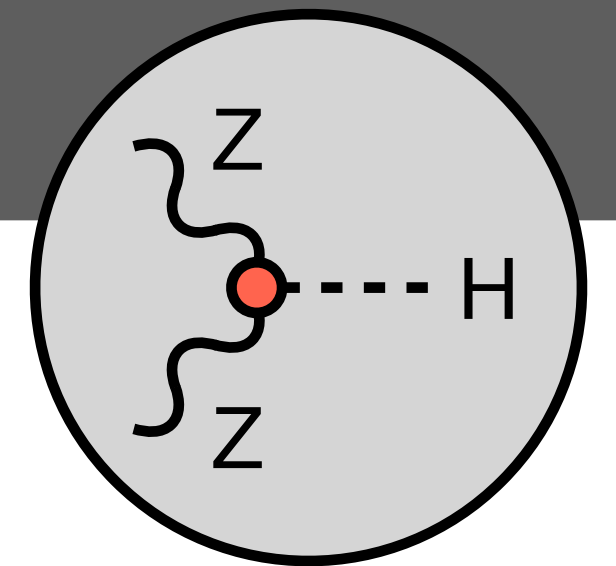
$$K_Z = -1$$

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!

Current CMS Result

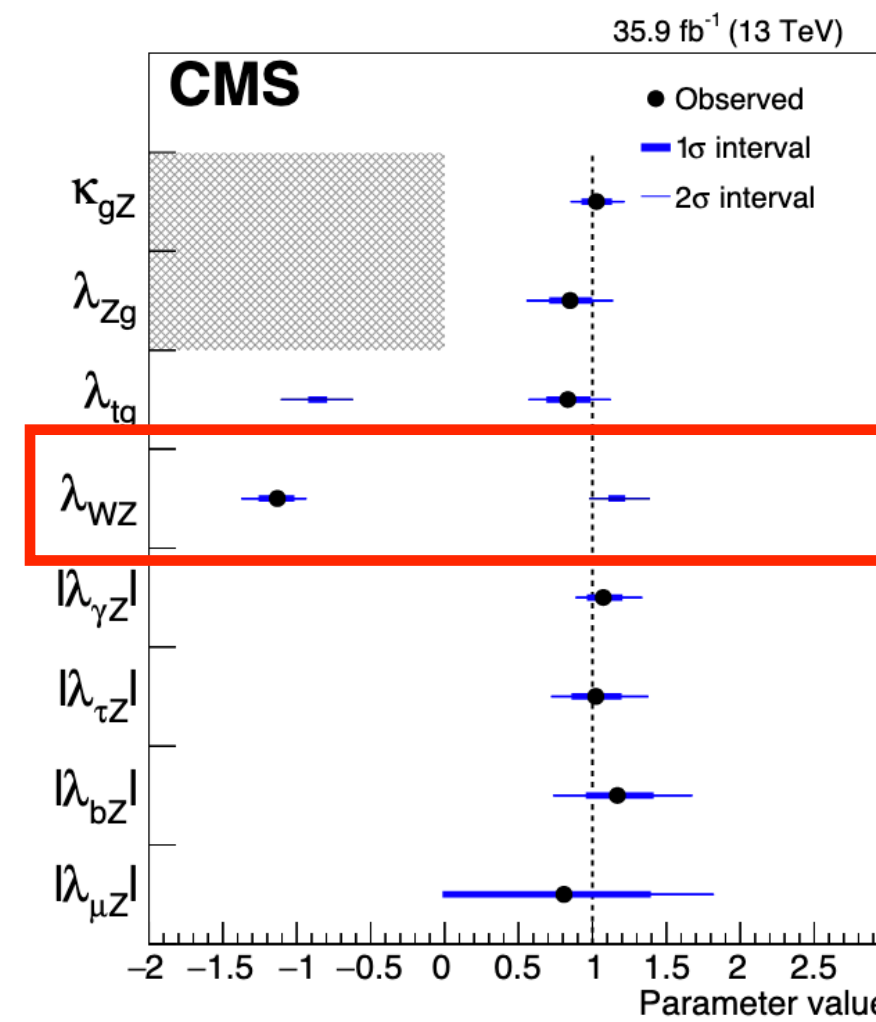
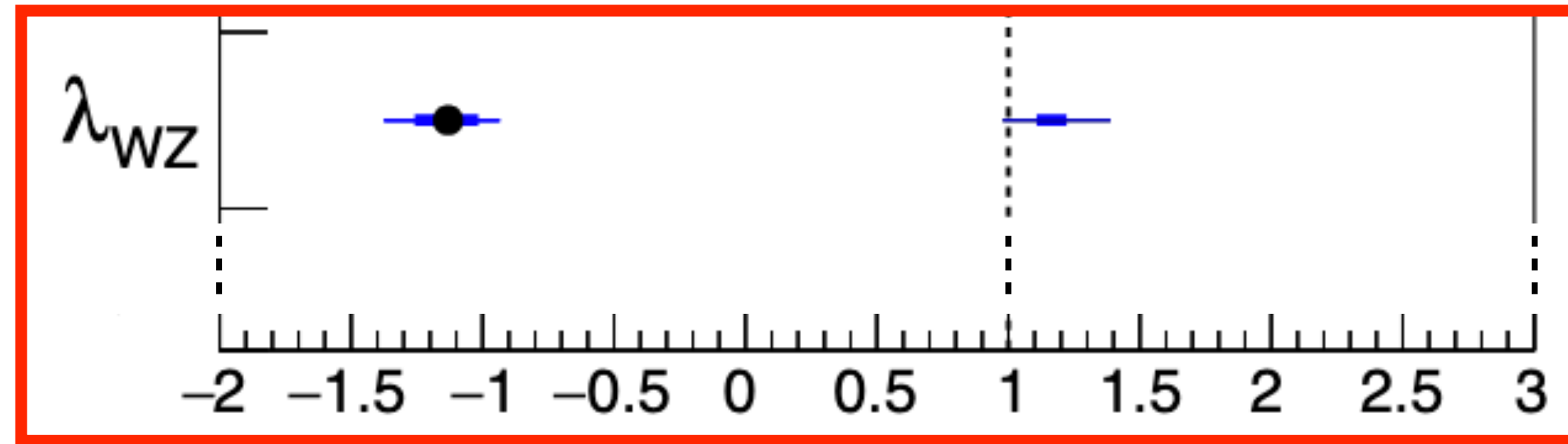


Figure 16: Summary of the model with coupling ratios and effective couplings for the ggH and $H \rightarrow \gamma\gamma$ loops. The points indicate the best fit values while the thick and thin horizontal bars show the 1σ and 2σ CL intervals, respectively. For this model, both positive and negative values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

Current CMS Result

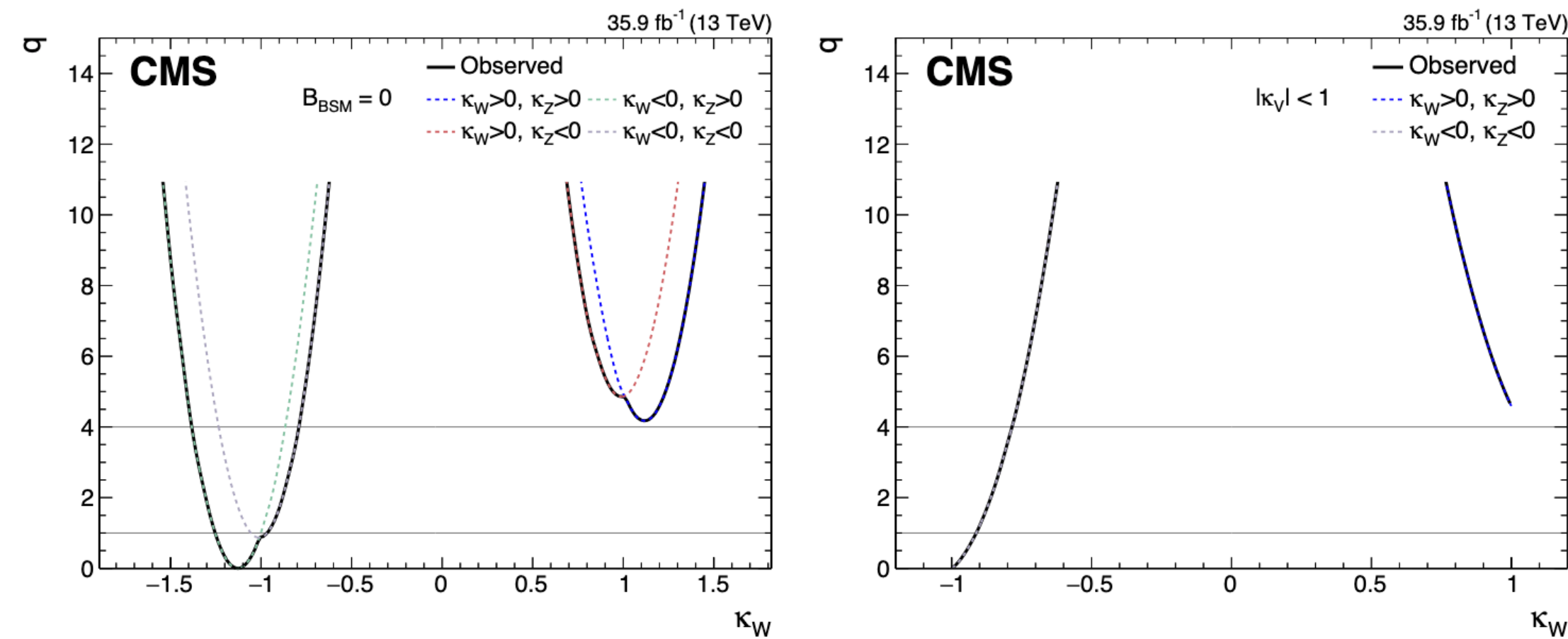


Figure 13: Scan of the test statistic q as a function of κ_W in the generic κ model assuming $B_{BSM} = 0$ (left) and allowing B_{inv} and B_{undet} to float (right). The different colored lines indicate the value of q for different combinations of signs for κ_W and κ_Z . The solid black line shows the 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.

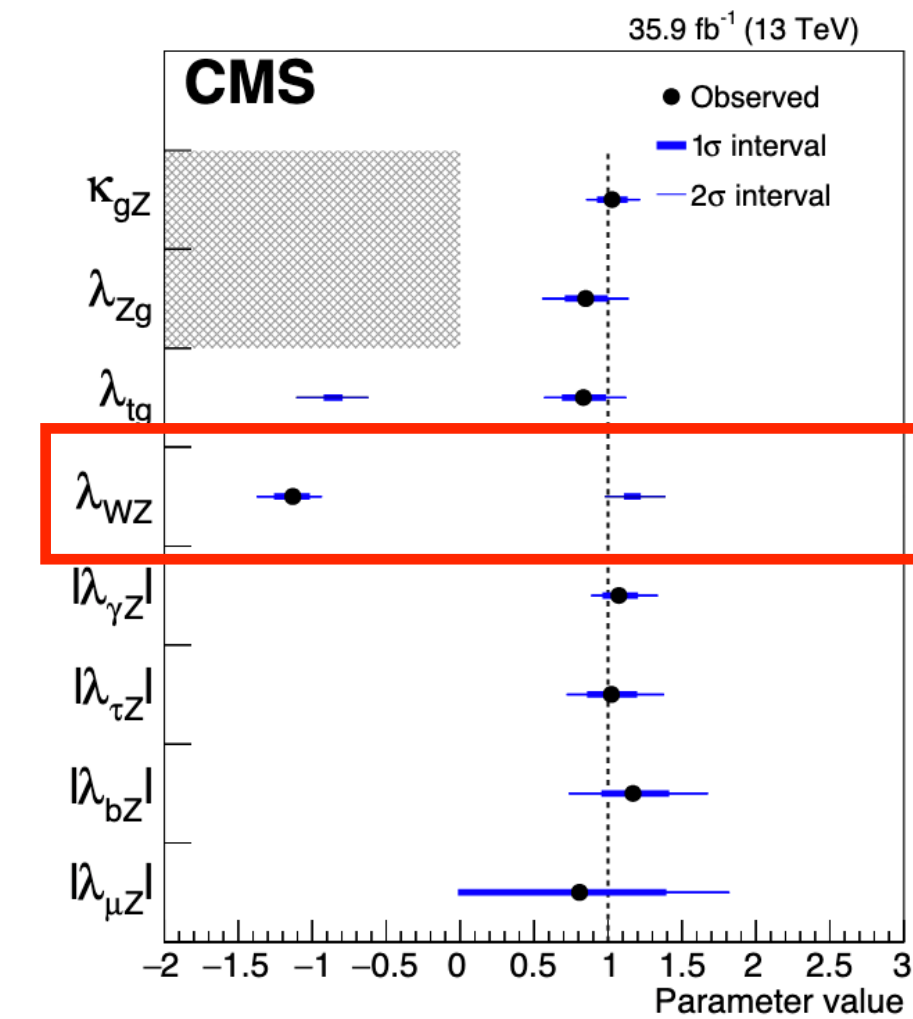
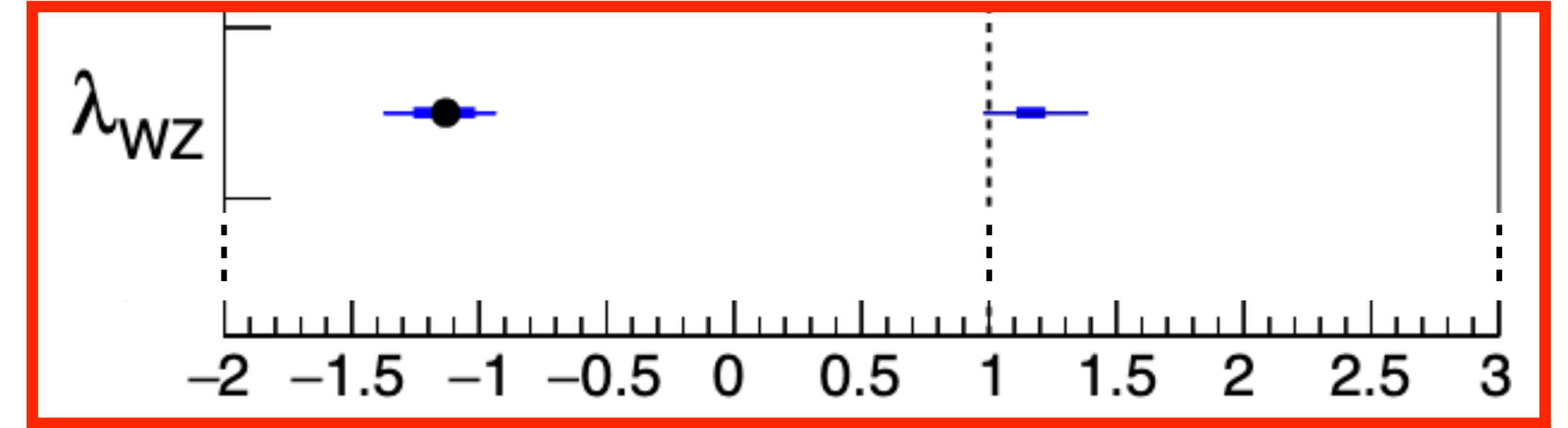


Figure 16: Summary of the model with coupling ratios and effective couplings for the ggH and $H \rightarrow \gamma\gamma$ loops. The points indicate the best fit values while the thick and thin horizontal bars show the 1σ and 2σ CL intervals, respectively. For this model, both positive and negative values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

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

VBS WH Cross Sections

Model	σ [pb]
$\kappa_W = \kappa_Z = +1$ (SM)	0.075
$\kappa_W = -1$	0.433
$\kappa_Z = -1$	0.433

↪ ×6

- Setting $\kappa_W = -1$ or $\kappa_Z = -1$ equivalently enhances cross section by a factor of 6
- 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 $\kappa_W = -1$** (kinematics are equivalent to $\kappa_Z = -1$)
 - Generated 100k UL NanoAOD events for 2016 pre-VFP, 2016 post-VFP, 2017, and 2018



Analysis Skim

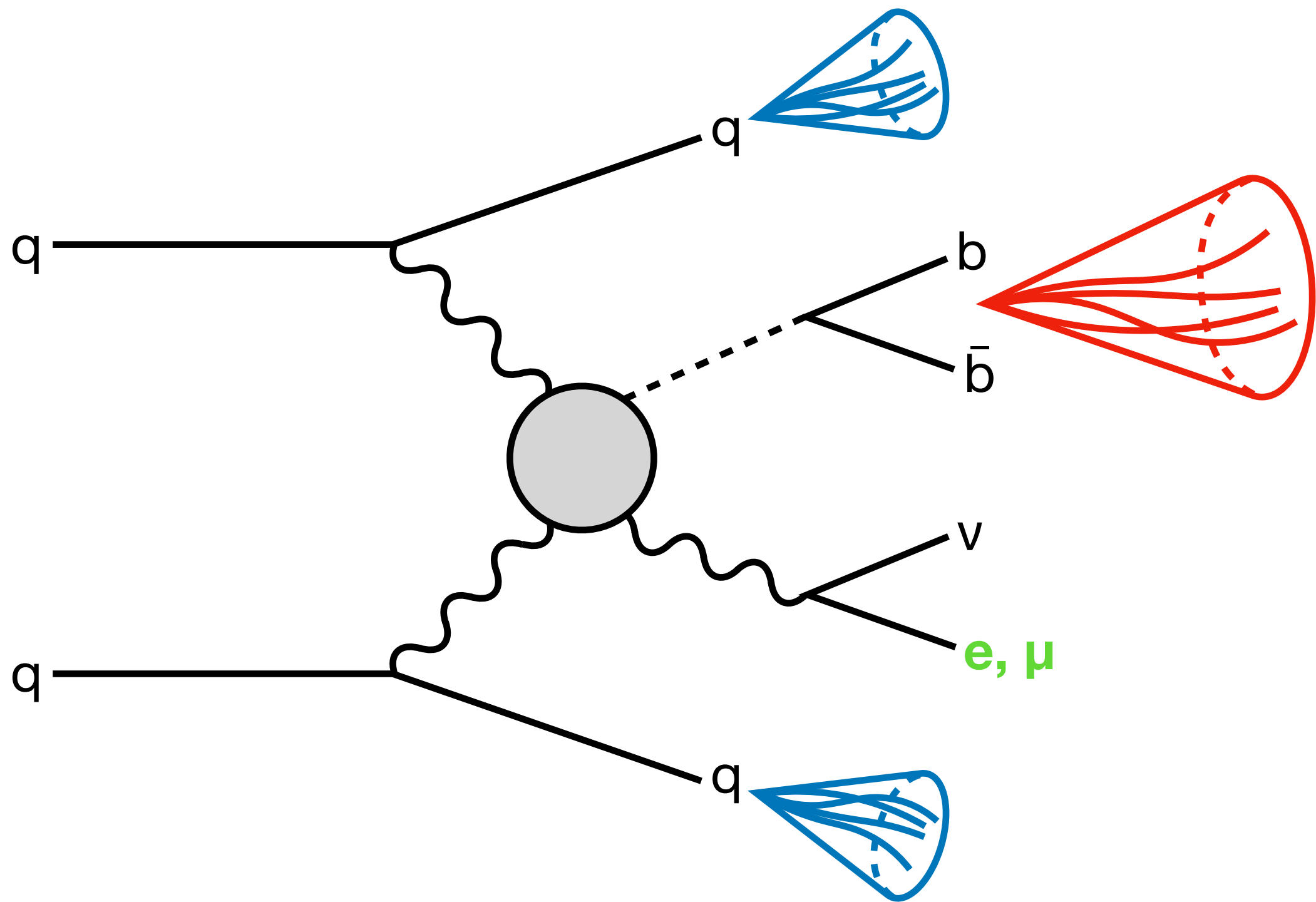
Yields scaled to $\text{lumi} \times \sigma$, rounded for readability

Cut	VH	VV/VVV/VBSWZ	W+Jets	SingleTop	TTbar+X	TTbar1L	TTbar2L	TotalBkg	Eff.*	VBSWH ($\kappa_W = -1$)	Eff.*
Skim AND Postskim	232	6.5K	122K	7.6K	1K	86K	10K	233K	—	2K	—

Object	Primary Skim	Postskim
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 $\Delta R(\text{fat jet, any veto lepton}) > 0.8$	≥ 1 w/ $p_T > 250$ GeV AND mass > 50 GeV AND $M_{SD} > 40$ GeV AND $\Delta R(\text{fat jet, tight lepton}) > 0.8$
Jets	≥ 1 w/ $p_T > 20$ GeV AND $\Delta R(\text{jet, any veto lepton}) > 0.4$	—
Other	—	$S_T > 800$ GeV

*Using the ttH lepton ID

VBS WH Objects



Object	Selections
One Lepton (loose veto in postskim)	• Passes ttH tight ID • $p_T > 40$ GeV
One Hbb Ak8 Jet	• $\Delta R(\text{Ak8 jet, any veto lep}) > 0.8$ • $p_T > 250$ GeV • $\text{mass} > 50$ GeV • $M_{SD}^* > 40$ GeV • Has max(PNet Xbb vs. QCD score) • PNet Xbb vs. QCD score > 0.3
Two VBS Ak4 jets	• $\Delta R(\text{Ak4 jet, any veto lep}) > 0.4$ • $\Delta R(\text{Ak4 jet, Hbb Ak8 jet}) > 0.8$ • $p_T > 30$ GeV • For > 2 candidates^{**}: • All in same η-hemisphere: leading/trailing (in P) jets • Else: leading (in P) jet from each η-hemisphere

*M_{SD} = soft drop mass
 **Hereafter referred to as the “Max-E” selection

Overview

Yields scaled to $\text{lumi} \times \sigma$, rounded for readability

Cut		VH	VV/VVV/VBSWZ	W+Jets	SingleTop	TTbar+X	TTbar1L	TTbar2L	TotalBkg	Eff.*	VBSWH ($\kappa_W = -1$)	Eff.*
Preselection	Skim AND Postskim	232	6.5K	122K	7.6K	1K	86K	10K	233K	—	2K	—
	$p_T(\ell) > 40 \text{ GeV}$	220	5.9K	110K	6.9K	910	78K	8.8K	211K	10%	1.8K	5%
	Passes single-lepton triggers	181	5.3K	98K	6.2K	773	68K	7.6K	186K	12%	1.6K	14%
	≥ 1 AK8 jet w/ $p_T > 250 \text{ GeV}$ AND mass $> 50 \text{ GeV}$ AND $M_{SD} > 40 \text{ GeV}$	181	5.3K	98K	6.2K	767	68K	7.4K	186K	0%	1.6K	0%
	2 VBS jets (max E)	50	2K	47K	4.4K	654	50K	6.3K	111K	40%	1.1K	30%
	$M_{jj} > 500 \text{ GeV}$ AND $ \Delta\eta_{jj} > 3$	7	376	7.7K	994	124	10K	1.4K	21K	81%	920	17%
	$\text{max(PNet Xbb score)} > 0.3$	3	51	587	537	77	5.8K	382	7.4K	64%	608	34%
	No b-tagged (med.) AK4 jets	2	43	484	194	20	1.7K	110	2.5K	88%	579	37%
SR	$M_{jj} > 1.5 \text{ TeV}$ AND $ \Delta\eta_{jj} > 4$	1	5	68	31	3	297	22	427	83%	322	44%
	$S_T > 1.2 \text{ TeV}$	0	1	11	6	0	25	2	45	89%	139	57%
	PNet Xbb vs. QCD > 0.9 AND $M_{SD} < 150$	0	0	1	0	0	2	1	4	91%	109	22%

- Using ParticleNet Xbb (mass-decorrelated) tagger
- Global b-veto for suppressing ttbar bkg: discard if there are *any* “good” b-tagged AK4s



*eff = 1 - after/before

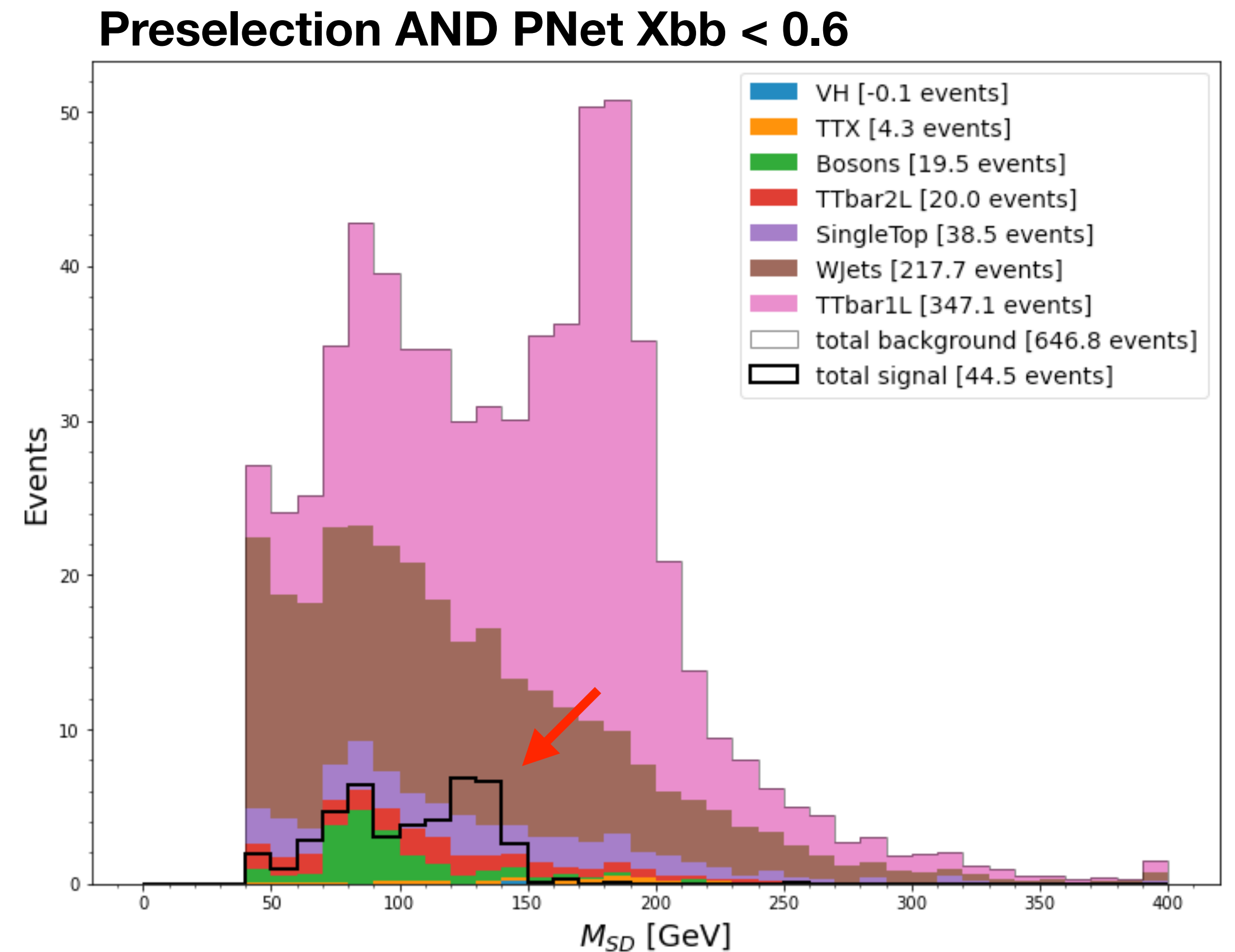
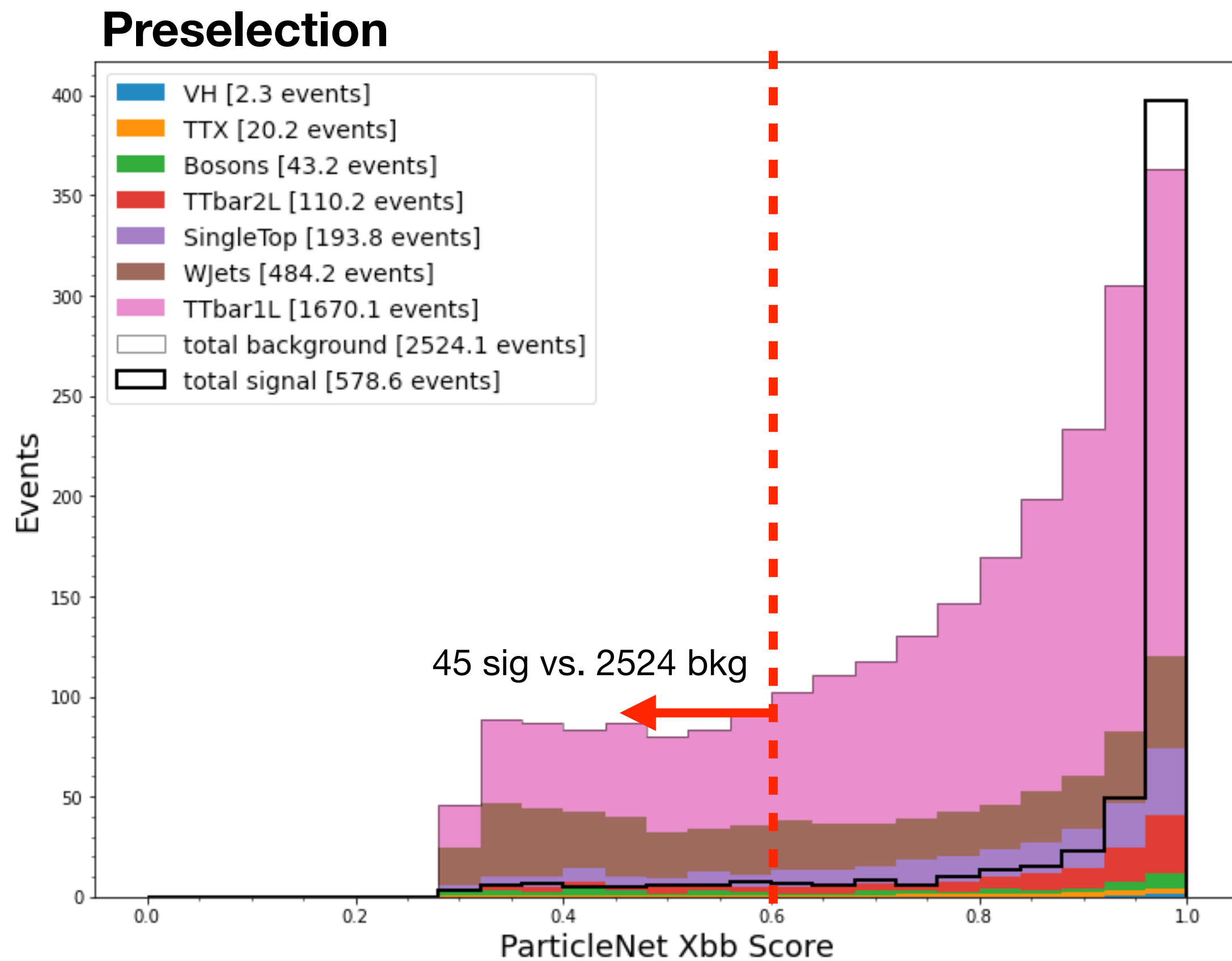


VBS WH Cutflow

Weighted yields (“wgt”) scaled to $lumi \times \sigma$

	VH		VV/VV/VBSWZ		W+Jets		SingleTop		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	raw	wgt
Bookkeeping	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8372794	233443.4	399990	59598.1
FindLeptons	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8372794	233443.4	399990	59598.1
Geq1VetoLep	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8372794	233443.4	129280	19148.8
Geq2Jets	—	—	Skim & Postskim already applied!										—	—	8372794	233443.4	128653	19057.0
Geq1FatJetNoVetoLepOverlap	—	—											—	—	8372794	233443.4	52537	7773.7
Exactly1TightLep	—	—											—	—	8372794	233443.4	31748	4734.4
Geq1FatJetNoTightLepOverlap	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8372794	233443.4	24525	3653.0
STgt800	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8372792	233443.3	13122	1950.0
SelectLeptons	63165	231.7	484277	6482.0	4603393	121610.6	76733	7558.5	939478	1042.8	1943369	86381.3	262377	10136.4	8372792	233443.3	13122	1950.0
Has1TightLep	63165	231.7	484277	6482.0	4603393	121610.6	76733	7558.5	939478	1042.8	1943369	86381.3	262377	10136.4	8372792	233443.3	13122	1950.0
LepPtGt40	60944	220.0	435912	5900.1	3889493	109822.4	68935	6915.5	817993	910.4	1761239	78242.7	227019	8754.4	7261535	210765.4	12416	1848.4
Passes1LepTriggers	54536	180.9	386803	5342.0	3325103	97557.5	61557	6223.5	700889	772.5	1535810	68294.5	197780	7636.6	6262478	186007.5	10736	1591.5
SelectFatJets	54536	180.9	386803	5342.0	3325103	97557.5	61557	6223.5	700889	772.5	1535810	68294.5	197780	7636.6	6262478	186007.5	10736	1591.5
Geq1FatJet	54470	181.1	384920	5337.1	3323420	97531.5	61404	6201.8	695035	767.0	1534845	68250.4	190755	7367.7	6244849	185636.6	10726	1590.3
SelectHbbFatJet	54470	181.1	384920	5337.1	3323420	97531.5	61404	6201.8	695035	767.0	1534845	68250.4	190755	7367.7	6244849	185636.6	10726	1590.3
SelectJetsNoHbbOverlap	54470	181.1	384920	5337.1	3323420	97531.5	61404	6201.8	695035	767.0	1534845	68250.4	190755	7367.7	6244849	185636.6	10726	1590.3
SaveAk4GlobalBVeto	54470	181.1	384920	5337.1	3323420	97531.5	61404	6201.8	695035	767.0	1534845	68250.4	190755	7367.7	6244849	185636.6	10726	1590.3
SelectVBSJetsMaxE	17662	50.3	251204	2035.9	2399176	47338.8	43921	4372.3	593328	654.0	1126440	50406.5	161098	6253.2	4592829	111110.9	7458	1113.0
MjjGt500_detajjGt3	3039	7.2	44926	376.4	393418	7744.0	10903	994.0	111929	124.2	231383	10337.4	35368	1371.5	830966	20954.6	6167	920.1
XbbGt0p3	2020	3.1	8633	50.8	34602	586.9	5486	537.0	71857	77.0	131324	5803.3	10113	382.3	264035	7440.3	4084	608.2
ApplyAk4GlobalBVeto	1627	2.4	6912	43.2	26699	484.2	1848	193.8	19216	20.2	37379	1670.1	2908	110.2	96589	2524.1	3877	578.6
MjjGt1500_detajjGt4	208	0.5	1127	4.6	5649	68.2	344	31.3	3046	3.5	6592	296.9	590	22.5	17556	427.5	2126	322.2
STGt1200	27	0.0	263	0.6	2244	11.3	56	6.1	500	0.5	565	24.7	49	1.8	3704	45.0	931	138.8
XbbScoreGt0p9_MSDLt150	23	0.0	65	0.0	217	1.3	1	0.2	44	0.0	45	2.0	13	0.5	408	4.0	729	108.8

Validation: Hbb Kinematics



PNet Xbb sideband not viable: visible signal bump in M_{SD}