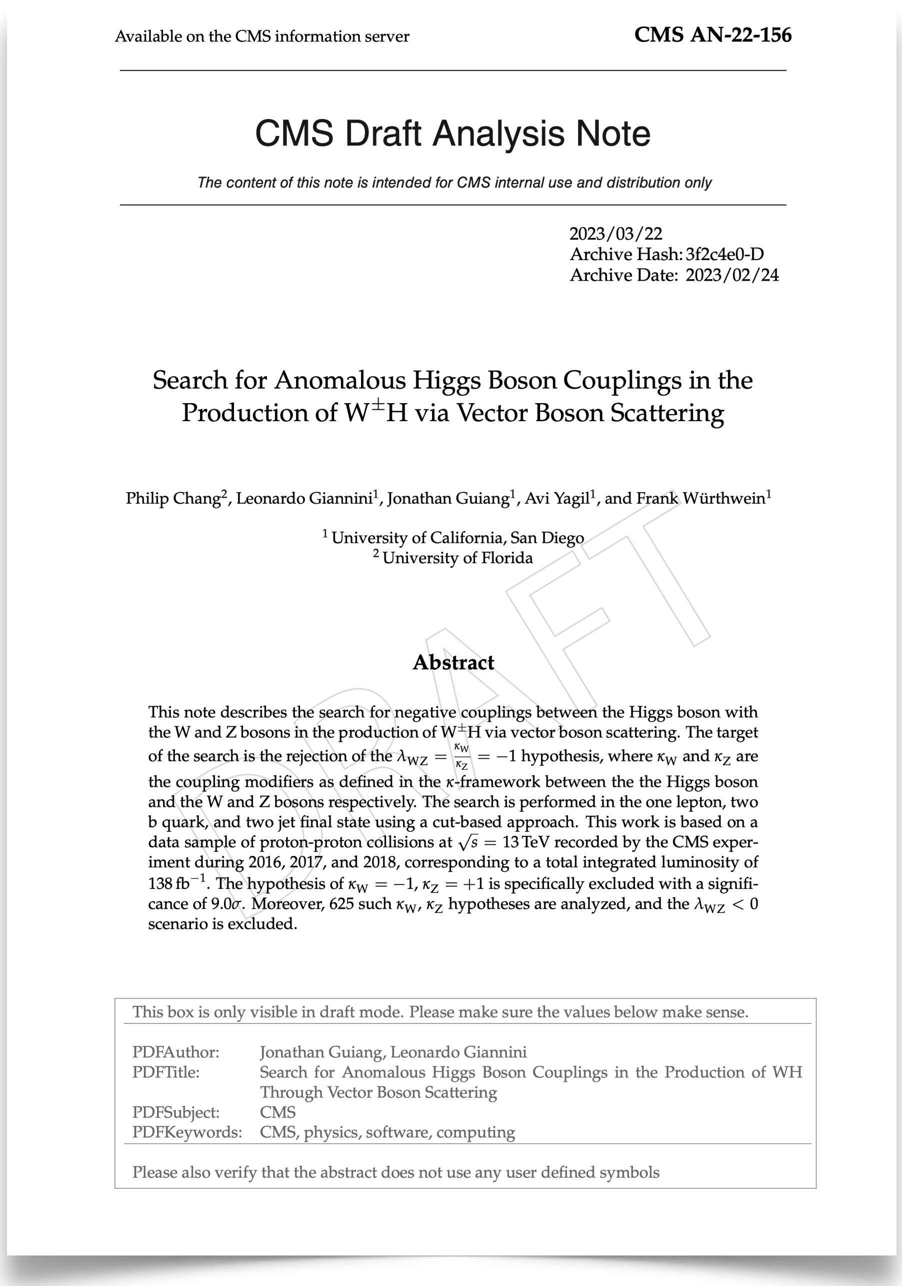
VBS WH Inaugural Talk

P. Chang[3], L. Giannini[2], **J. Guiang**[2], A. Levin[1], Q. Li[1], S. Qian[1], F. Würthwein[2]

[1]Peking University
[2]University of California, San Diego
[3]University of Florida

CMS Weekly General Meeting

Documentation



Available on the CMS information server **CMS AN-22-156**

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2023/03/22
Archive Hash: 3f2c4e0-D
Archive Date: 2023/02/24

Search for Anomalous Higgs Boson Couplings in the Production of $W^{\pm}H$ via Vector Boson Scattering

Philip Chang[2], Leonardo Giannini[1], Jonathan Guiang[1], Avi Yagil[1], and Frank Würthwein[1]

[1] University of California, San Diego
[2] University of Florida

Abstract

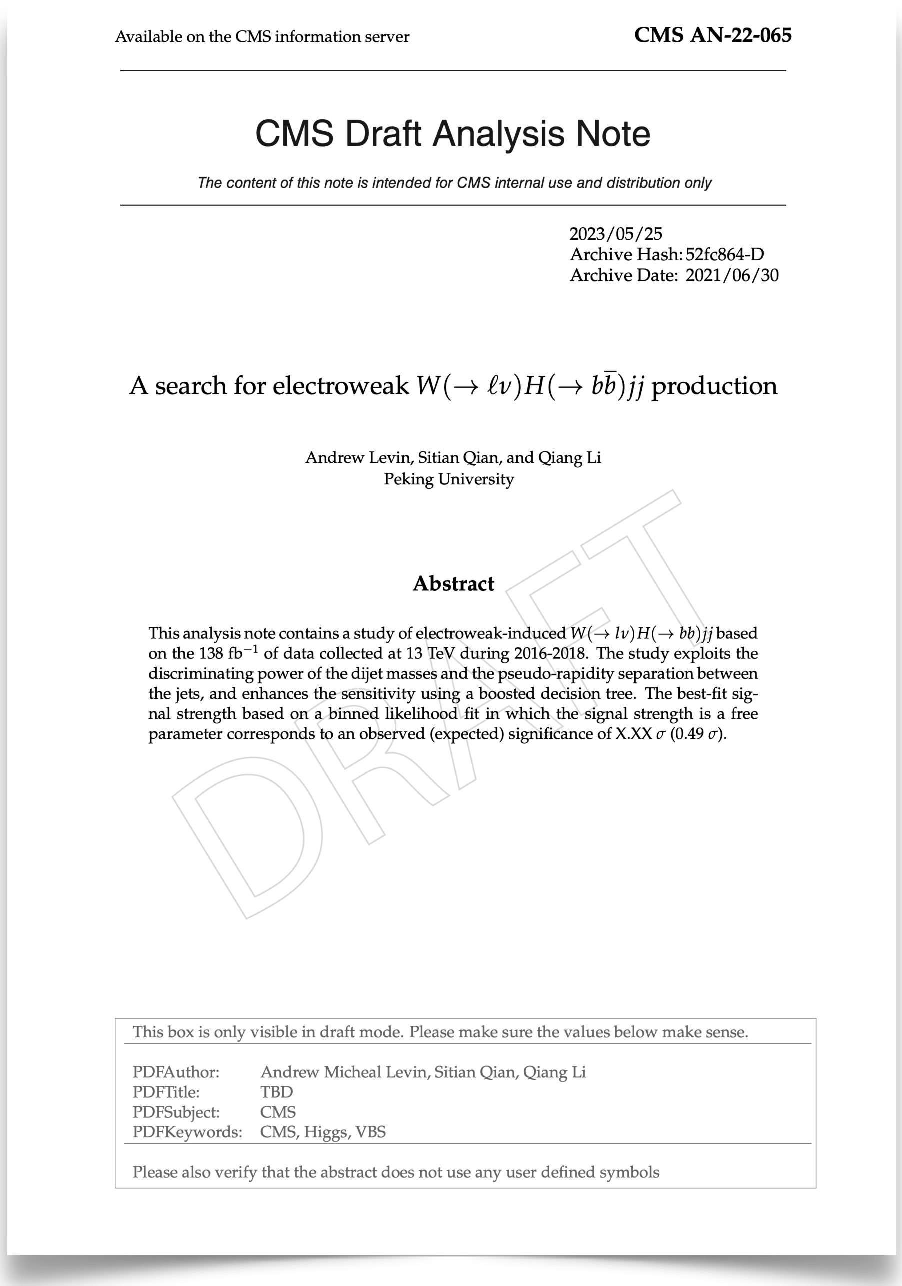
This note describes the search for negative couplings between the Higgs boson with the W and Z bosons in the production of $W^{\pm}H$ via vector boson scattering. The target of the search is the rejection of the $\lambda_{WZ} = \frac{\kappa_W}{\kappa_Z} = -1$ hypothesis, where κ_W and κ_Z are the coupling modifiers as defined in the κ-framework between the the Higgs boson and the W and Z bosons respectively. The search is performed in the one lepton, two b quark, and two jet final state using a cut-based approach. This work is based on a data sample of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded by the CMS experiment during 2016, 2017, and 2018, corresponding to a total integrated luminosity of 138 fb^{-1}. The hypothesis of $\kappa_W = -1$, $\kappa_Z = +1$ is specifically excluded with a significance of 9.0σ. Moreover, 625 such κ_W, κ_Z hypotheses are analyzed, and the $\lambda_{WZ} < 0$ scenario is excluded.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Jonathan Guiang, Leonardo Giannini
PDFTitle:	Search for Anomalous Higgs Boson Couplings in the Production of WH Through Vector Boson Scattering
PDFSubject:	CMS
PDFKeywords:	CMS, physics, software, computing

Please also verify that the abstract does not use any user defined symbols

+



Available on the CMS information server **CMS AN-22-065**

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2023/05/25
Archive Hash: 52fc864-D
Archive Date: 2021/06/30

A search for electroweak $W(\rightarrow \ell\nu)H(\rightarrow b\bar{b})jj$ production

Andrew Levin, Sitian Qian, and Qiang Li
Peking University

Abstract

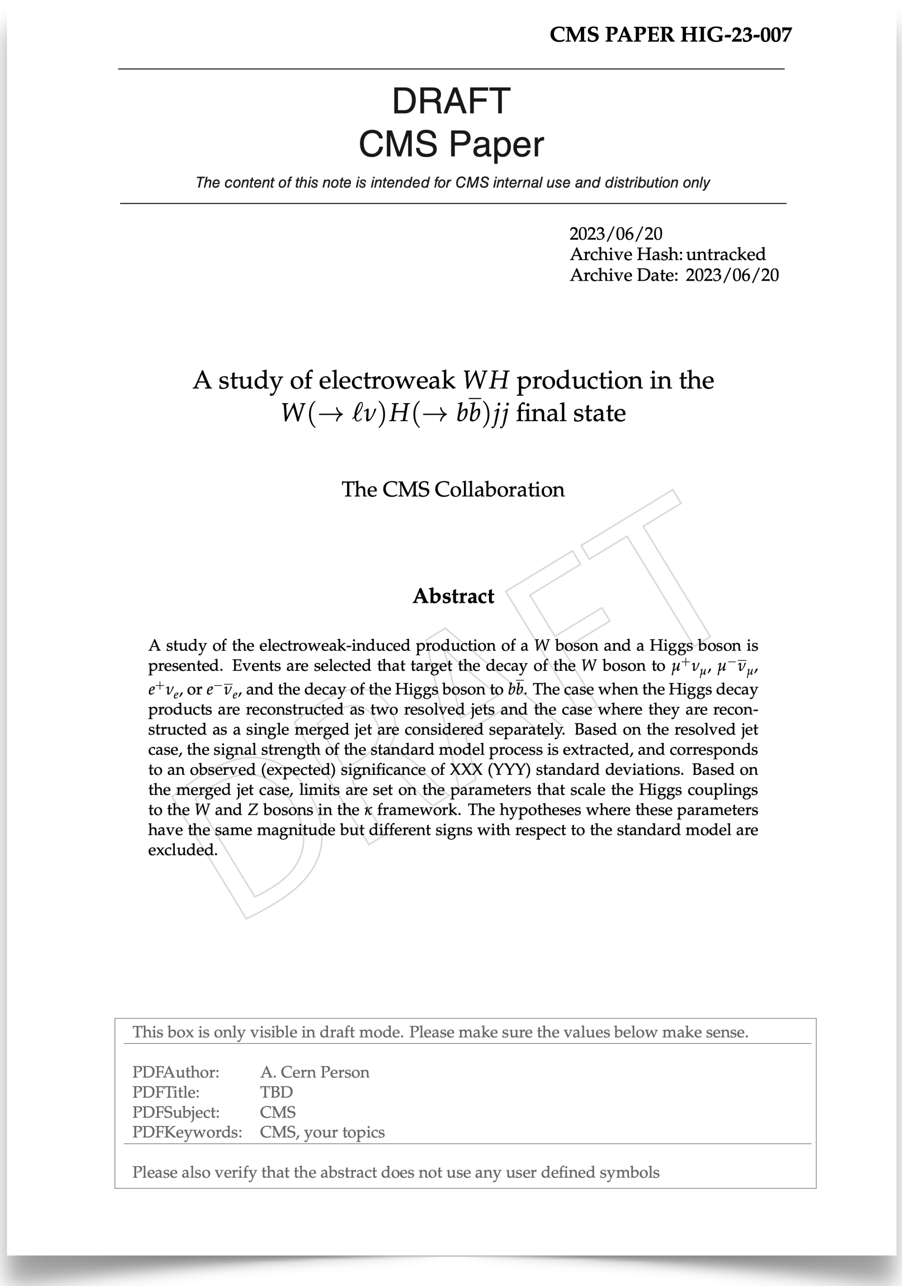
This analysis note contains a study of electroweak-induced $W(\rightarrow l\nu)H(\rightarrow bb)jj$ based on the 138 fb^{-1} of data collected at 13 TeV during 2016-2018. The study exploits the discriminating power of the dijet masses and the pseudo-rapidity separation between the jets, and enhances the sensitivity using a boosted decision tree. The best-fit signal strength based on a binned likelihood fit in which the signal strength is a free parameter corresponds to an observed (expected) significance of X.XX σ (0.49 σ).

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Andrew Micheal Levin, Sitian Qian, Qiang Li
PDFTitle:	TBD
PDFSubject:	CMS
PDFKeywords:	CMS, Higgs, VBS

Please also verify that the abstract does not use any user defined symbols

=



CMS PAPER HIG-23-007

DRAFT
CMS Paper

The content of this note is intended for CMS internal use and distribution only

2023/06/20
Archive Hash: untracked
Archive Date: 2023/06/20

A study of electroweak WH production in the $W(\rightarrow \ell\nu)H(\rightarrow b\bar{b})jj$ final state

The CMS Collaboration

Abstract

A study of the electroweak-induced production of a W boson and a Higgs boson is presented. Events are selected that target the decay of the W boson to $\mu^+\nu_\mu$, $\mu^-\bar{\nu}_\mu$, $e^+\nu_e$, or $e^-\bar{\nu}_e$, and the decay of the Higgs boson to $b\bar{b}$. The case when the Higgs decay products are reconstructed as two resolved jets and the case where they are reconstructed as a single merged jet are considered separately. Based on the resolved jet case, the signal strength of the standard model process is extracted, and corresponds to an observed (expected) significance of XXX (YYY) standard deviations. Based on the merged jet case, limits are set on the parameters that scale the Higgs couplings to the W and Z bosons in the κ framework. The hypotheses where these parameters have the same magnitude but different signs with respect to the standard model are excluded.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	A. Cern Person
PDFTitle:	TBD
PDFSubject:	CMS
PDFKeywords:	CMS, your topics

Please also verify that the abstract does not use any user defined symbols

AN-2022/156
(boosted)

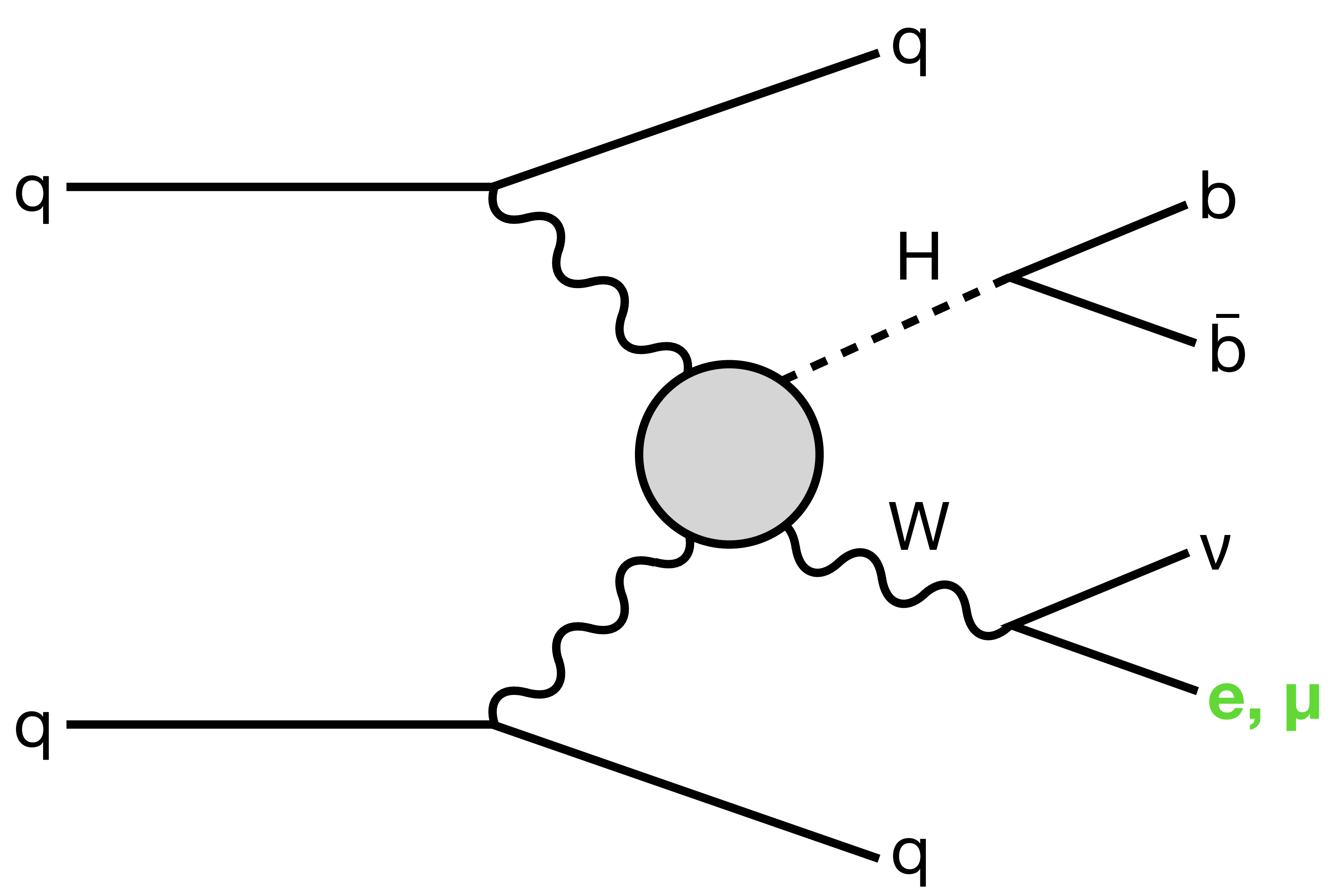
AN-2022/065
(resolved)

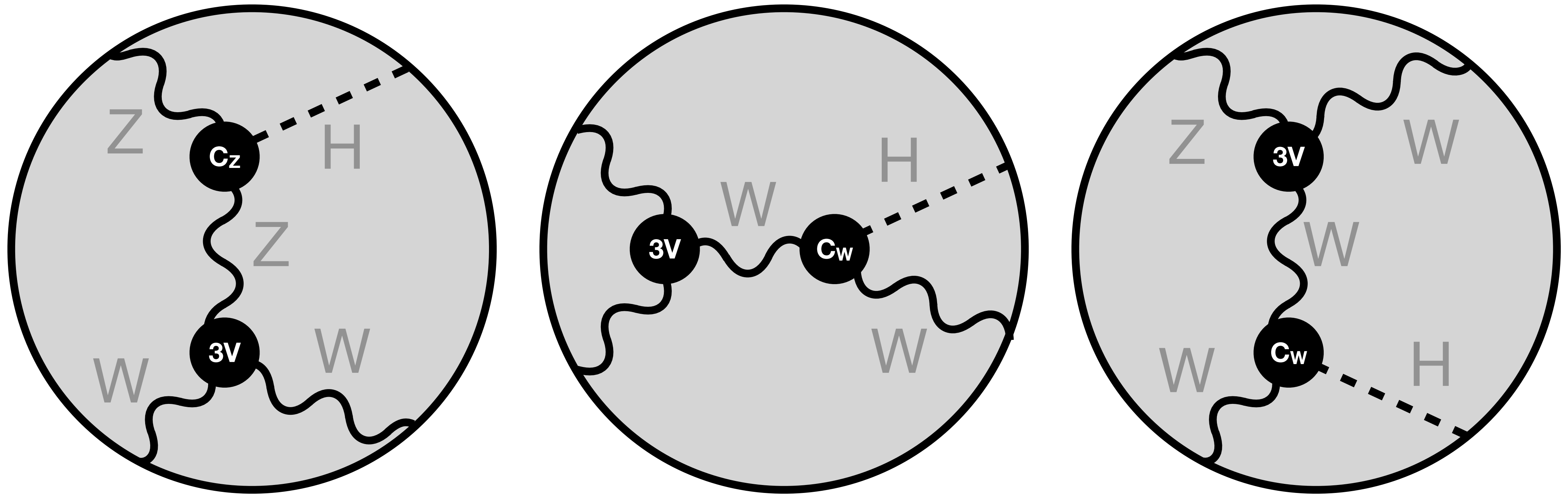
HIG-23-007

PubTalk: https://cms-pub-talk.web.cern.ch/c/hig/hig-23-007
CADI: https://cms.cern.ch/iCMS/analysisadmin/cadilines?line=HIG-23-007
TWiki: https://twiki.cern.ch/twiki/bin/viewauth/CMS/ElectroweakWHjjQA

Overview

- Targeting **VBS WH**
  - In particular: $H \to b\bar{b}$ and $W \to \ell\nu$
- **Boosted** analysis (UCSD + UFL)
  - $H \to b\bar{b}$ reconstructed as a single AK8 jet
  - Targeting an exclusion of BSM κ_W/κ_Z values
- **Resolved** analysis (PKU)
  - $H \to b\bar{b}$ reconstructed as two AK4 jets
  - Targeting an observation of SM VBS WH

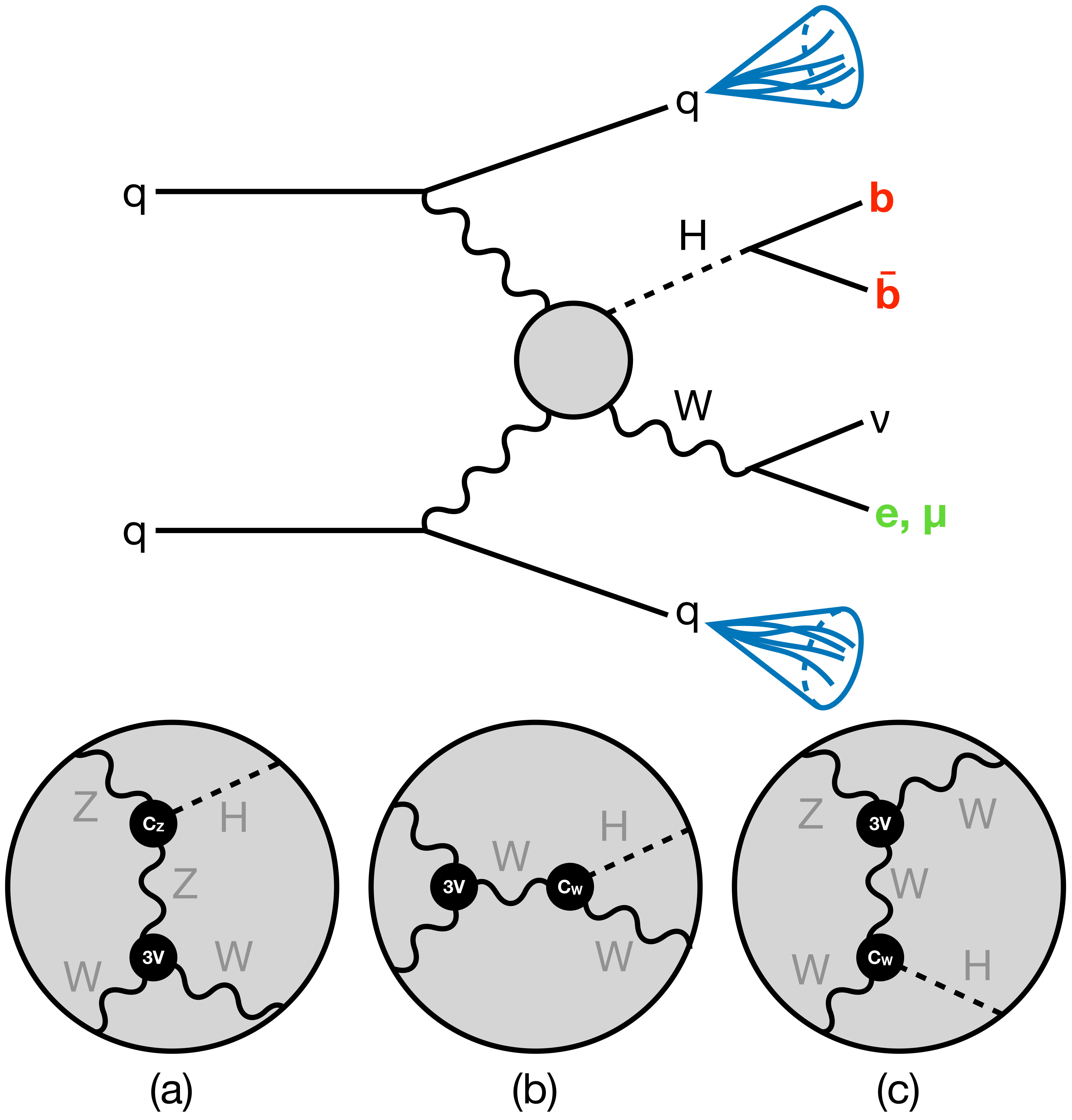




(a) (b) (c)

Signal Signature

- **VBS** WH→ℓν$b\bar{b}$ signature:
  - **VBS quarks** → 2 jets w/ large $\Delta\eta_{jj}$, M_{jj}
  - **H→$b\bar{b}$**
    - Most favorable BR
    - **Boosted:** 1 fat jet tagged w/ **ParticleNet**
    - **Resolved:** 2 jets tagged w/ **DeepJet**
  - One and only **one lepton**
    - Used for trigger/cleaner signature



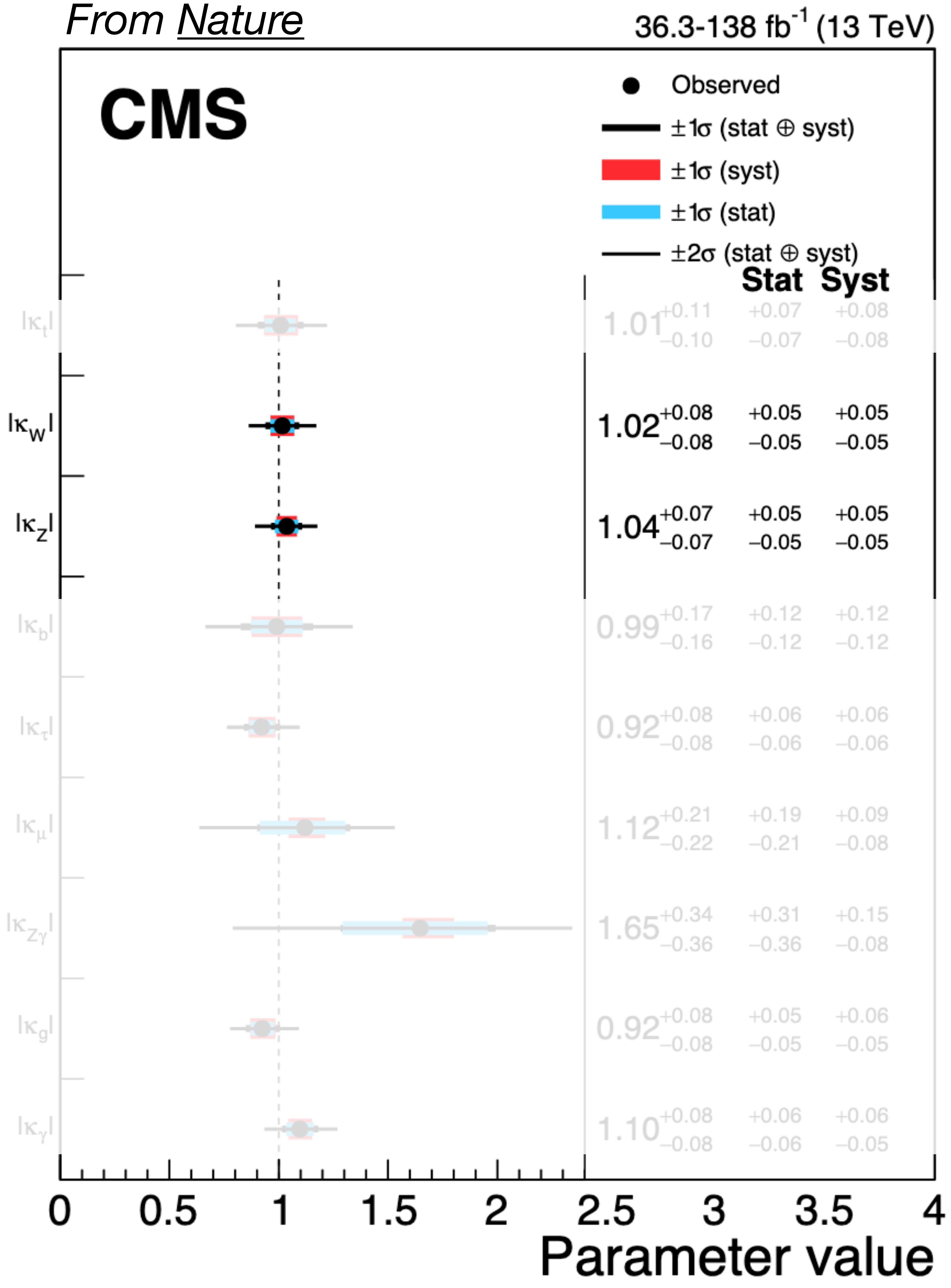
UC San Diego

Boosted Analysis

Target Higgs Couplings

- CMS has already pinned **$|\kappa_W| = 1$** and **$|\kappa_Z| = 1$**
    - Within an **uncertainty of 10%**
    - Effectively restricted to κ_V^2, so only know **magnitude**
- SM predicts that they are the same sign
    - i.e. we **expect $\lambda_{WZ} = \kappa_W/\kappa_Z = +1$**
- We have thus far not confirmed this prediction
    - Fun fact: best CMS limit* slightly prefers λ_{WZ} = -1
- **Need a process that is linear in κ_V**

From Nature



*https://link.springer.com/article/10.1140/epjc/s10052-019-6909-y

Enter: VBS WH

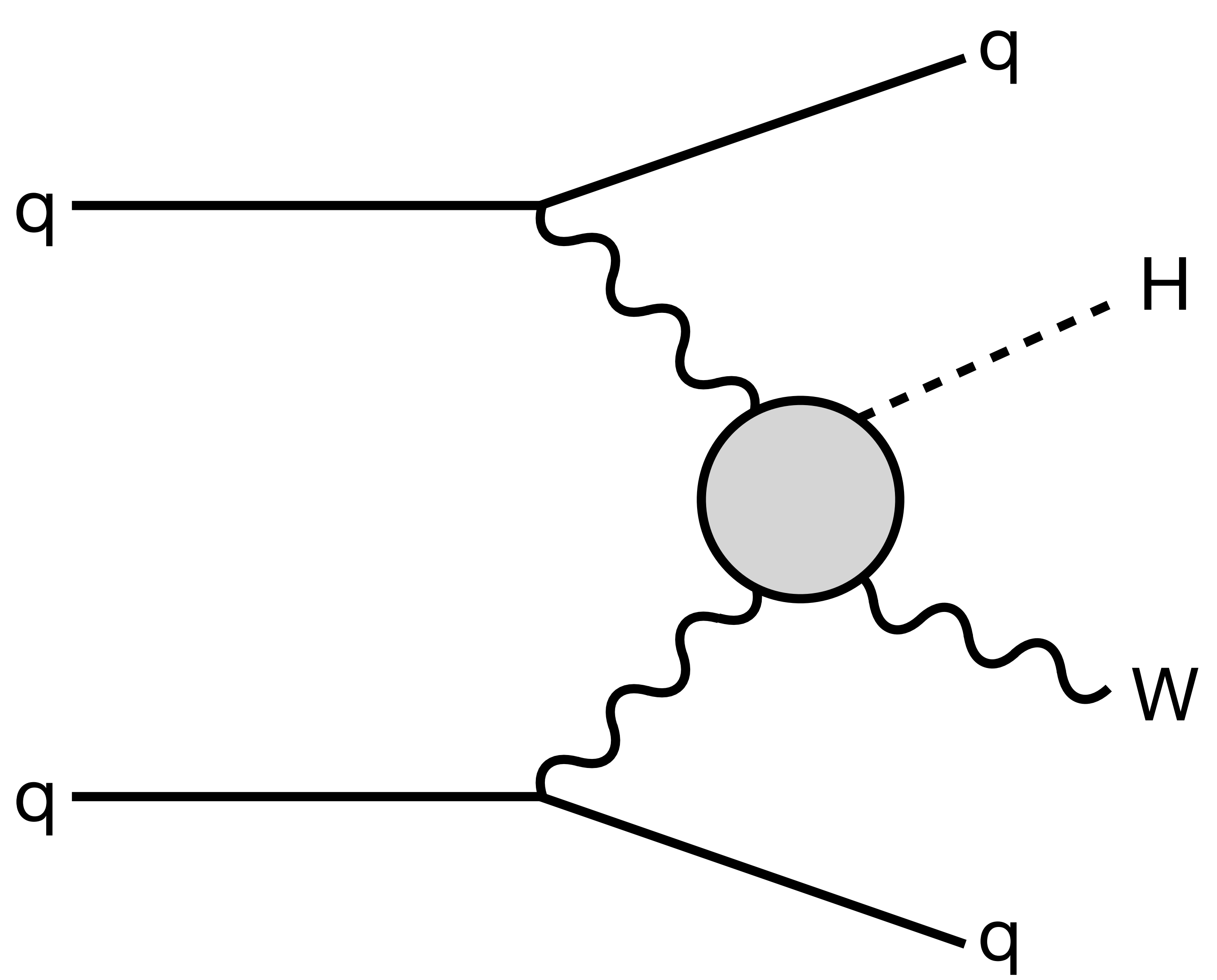
- Targeting **VBS WH**
  - σ = 0.075 pb (λ_{WZ} = +1)
  - σ = 0.433 pb (λ_{WZ} = -1)

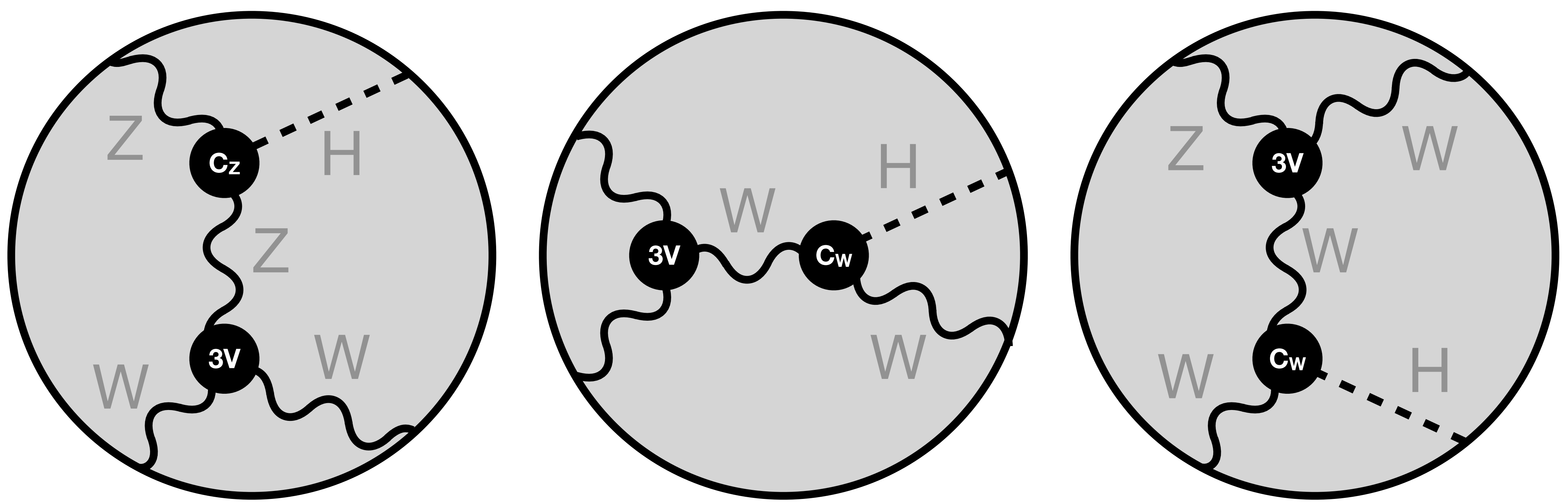
×6

- **Linear in κ_V**

$$\sigma \propto |\mathscr{M}|^2 = \kappa_W^2 |\mathscr{M}_W|^2 + \kappa_W \kappa_Z \mathscr{M}_{WZ}^2 + \kappa_Z^2 |\mathscr{M}_Z|^2$$

- We present an analysis that can **strongly exclude** the $\lambda_{WZ} < 0$ scenario
- Optimized for λ_{WZ} = -1



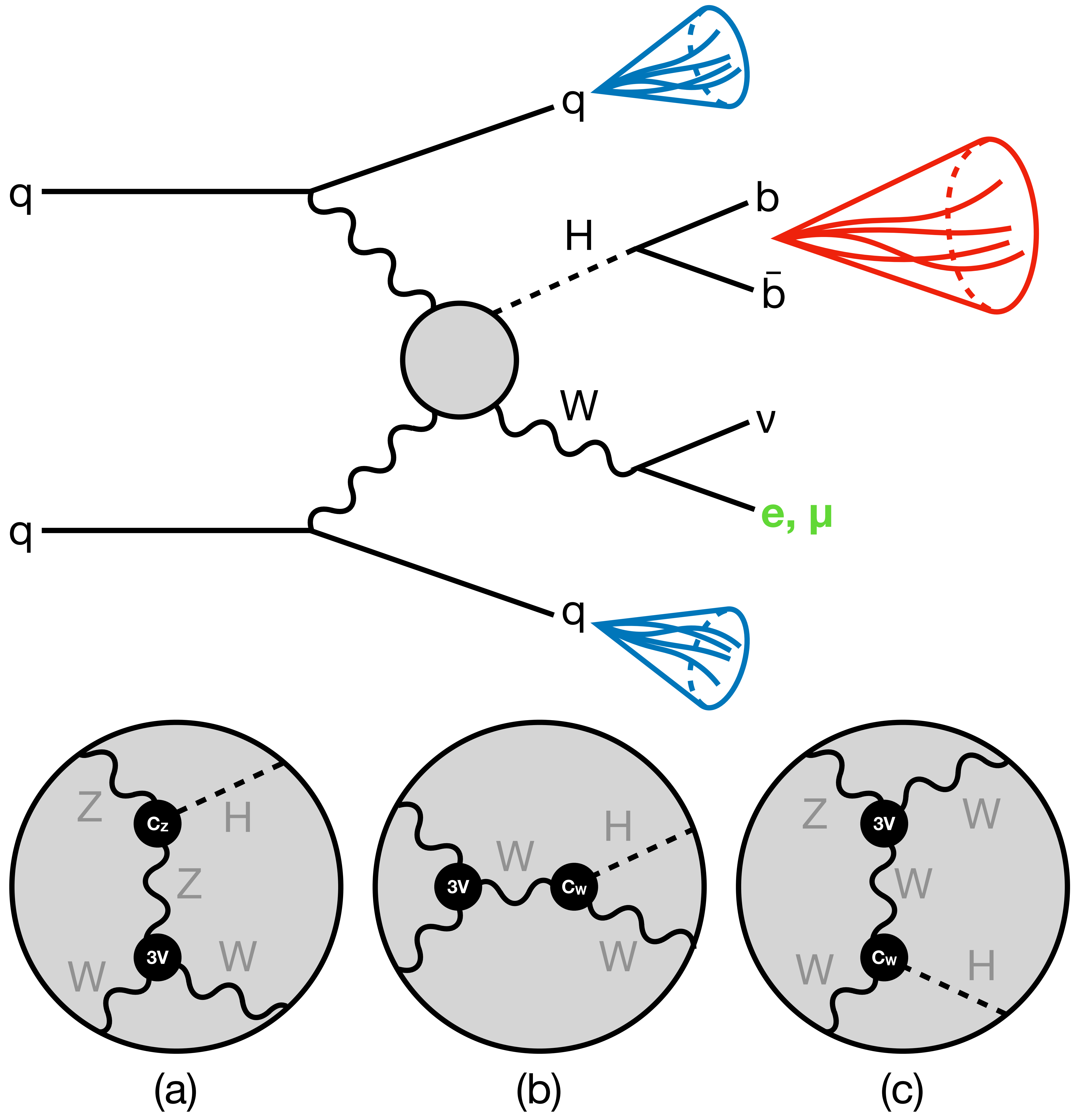


(a) (b) (c)

UC San Diego

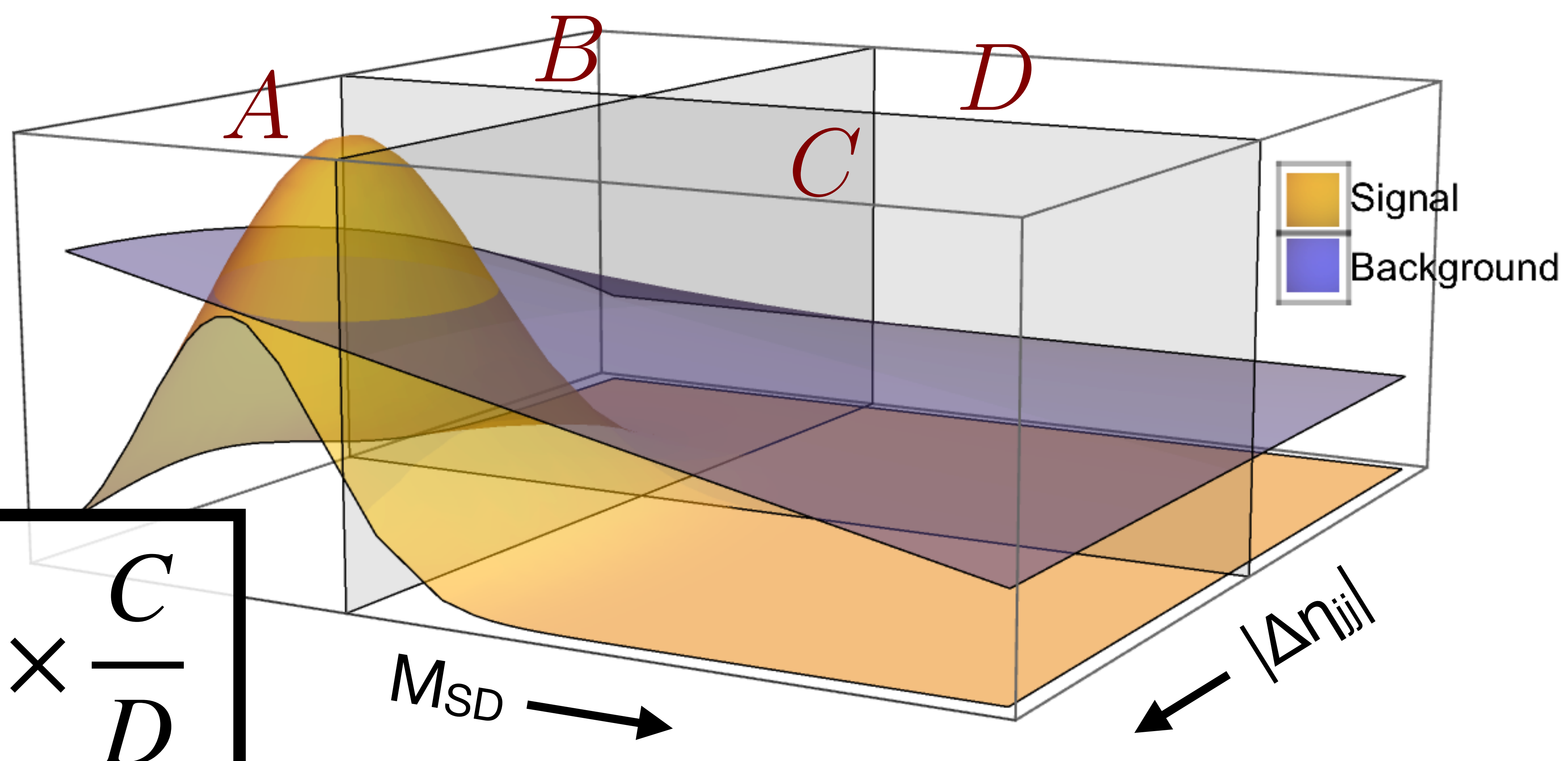
Analysis Strategy

1. Leverage High Level Triggers (HLTs) for most basic selection
   - We use the **single lepton** triggers
2. Construct a Signal Region (SR) with a large signal-to-background ratio
   - Roughly 370 sig. vs. 120 bkg.
3. Implement a data-driven estimation of the background in the SR
4. Perform a **simple counting experiment**

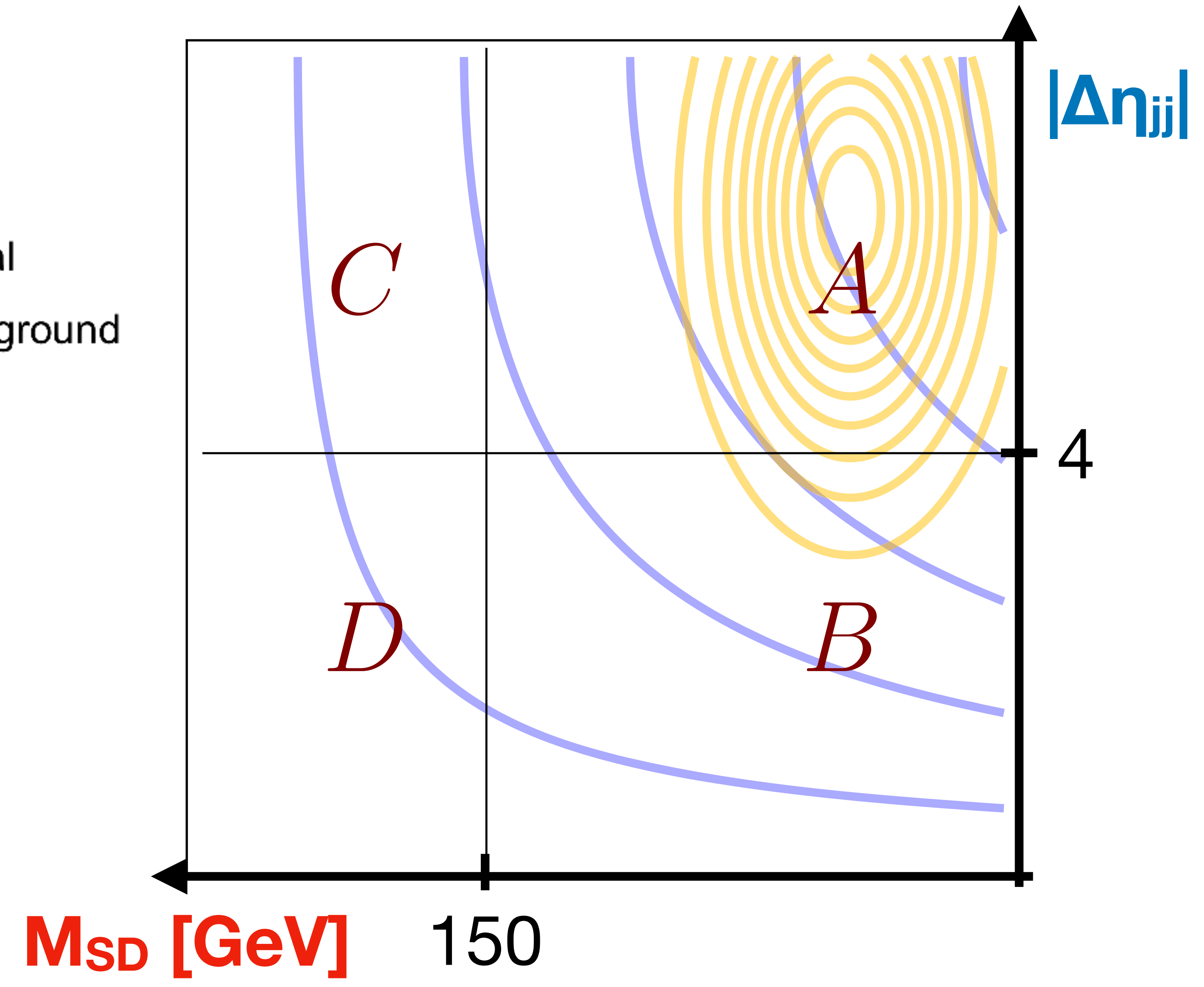


Data-Driven Background Extrapolation

Diagram from G. Kasieczka et al.



$$A^{pred} = B \times \frac{C}{D}$$

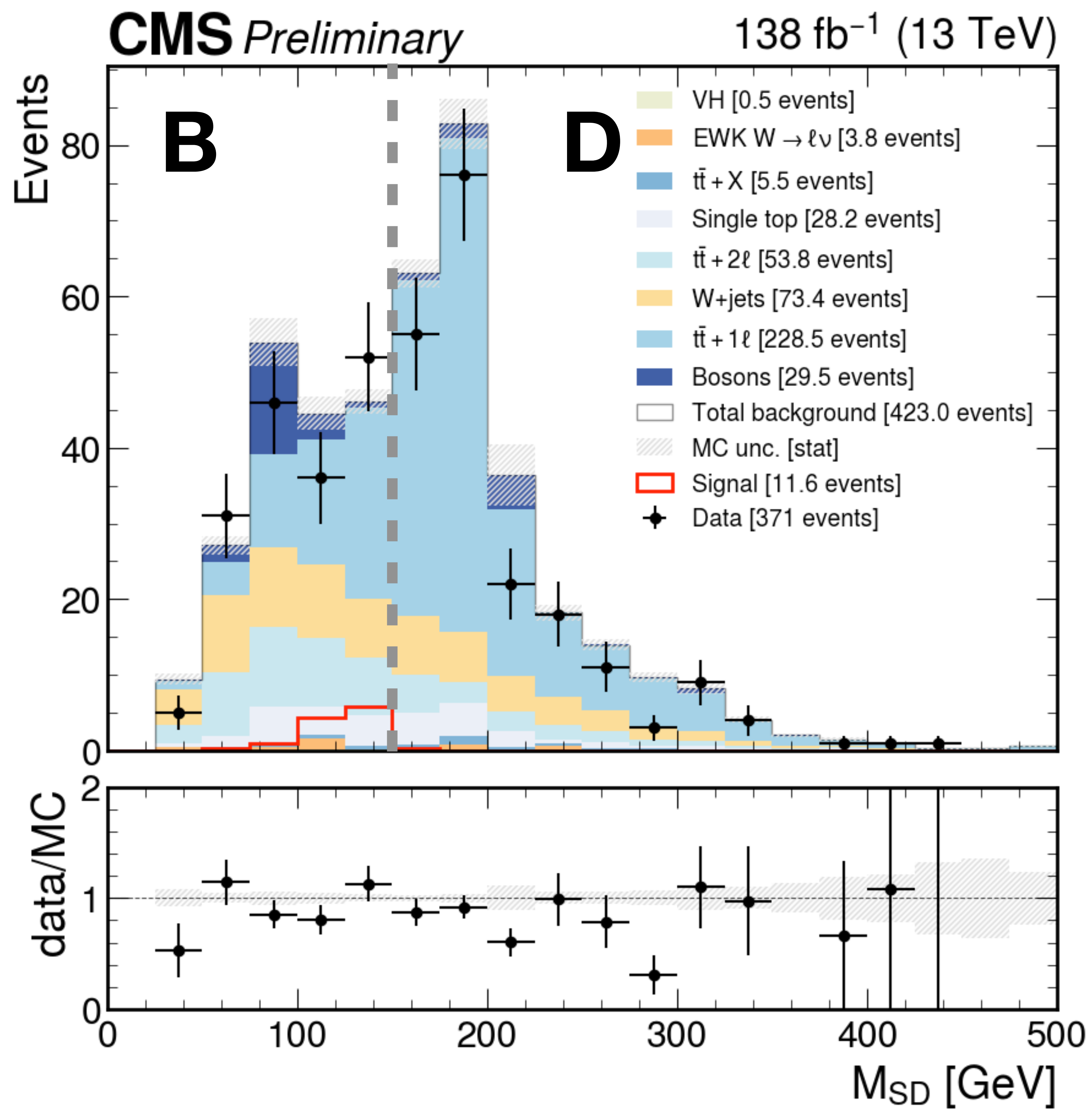


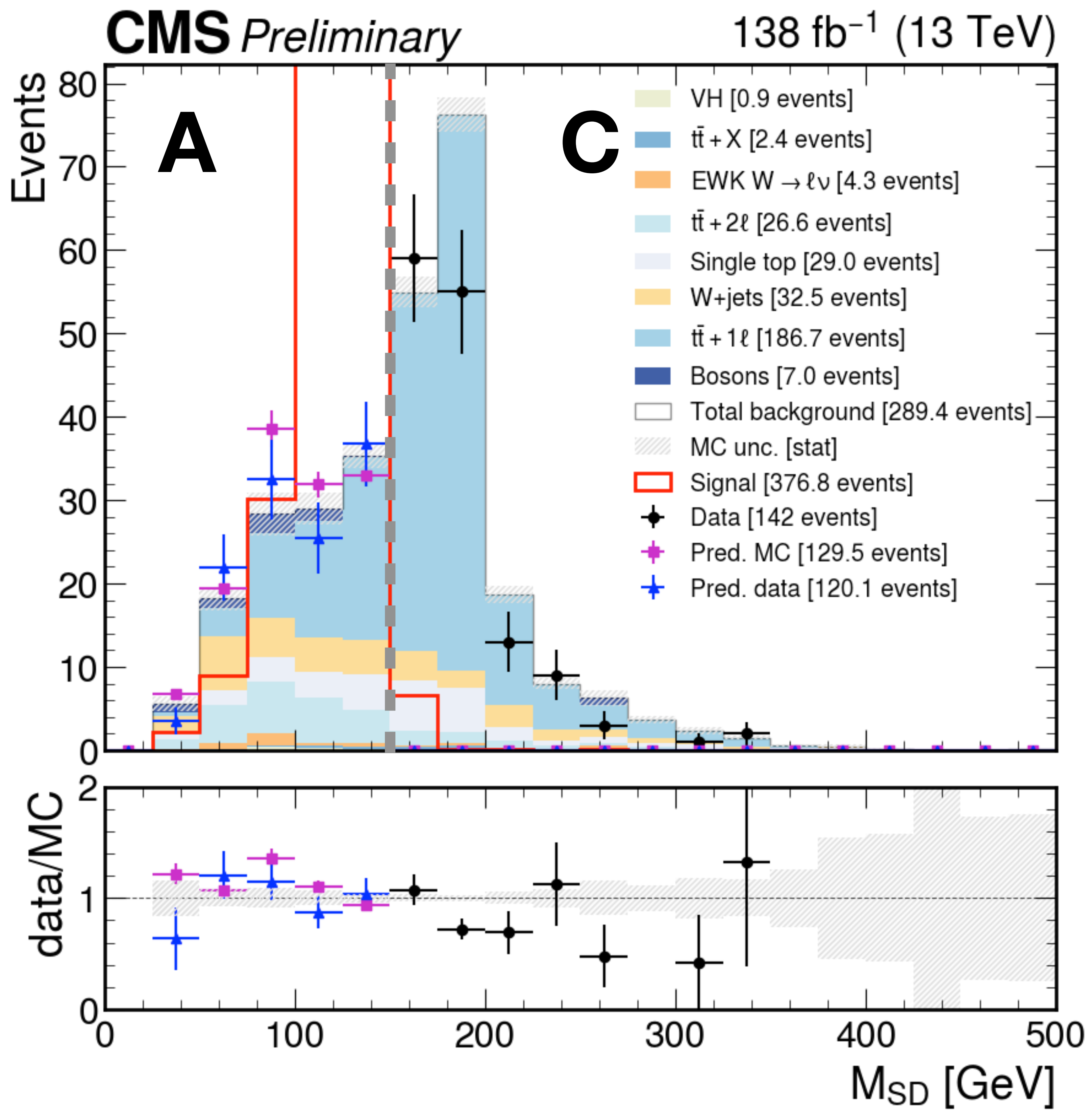
- Signal region* (Region A):
  - M_{jj} > 600 GeV & **$|\Delta\eta_{jj}|$ > 4** & **M_{SD} < 150 GeV** & ParticleNet Xbb > 0.9 & S_T > 900 GeV
- Background is predominantly from $t\bar{t}$+1ℓ production
- We use the ABCD ($|\Delta\eta_{jj}|$ vs. M_{jj}) method as above to estimate all bkg.

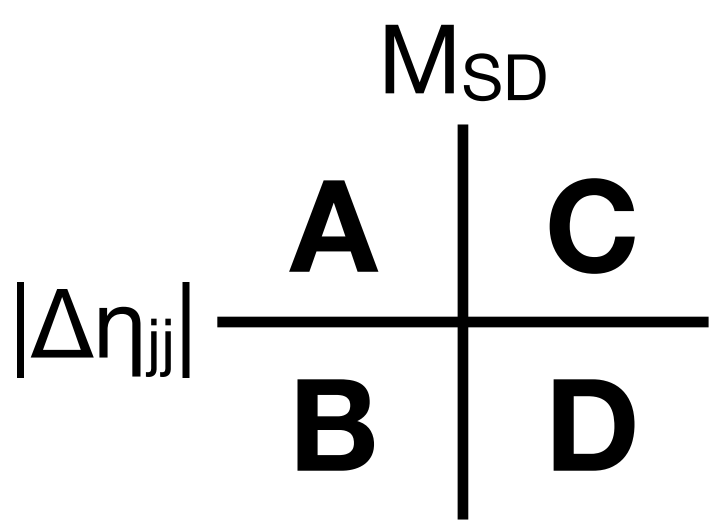
*Defined within Preselection region (detailed in backup)

UC San Diego

Data-Driven Background Extrapolation







$$A^{pred}_{MC} = B \times \frac{C_{MC}}{D_{MC}} = 129 \text{ events}$$

Compared to actual MC yield ⇒ **method closes well**

$$A^{pred}_{data} = B \times \frac{C_{data}}{D_{data}} = 120 \text{ events}$$

Used for **final prediction**

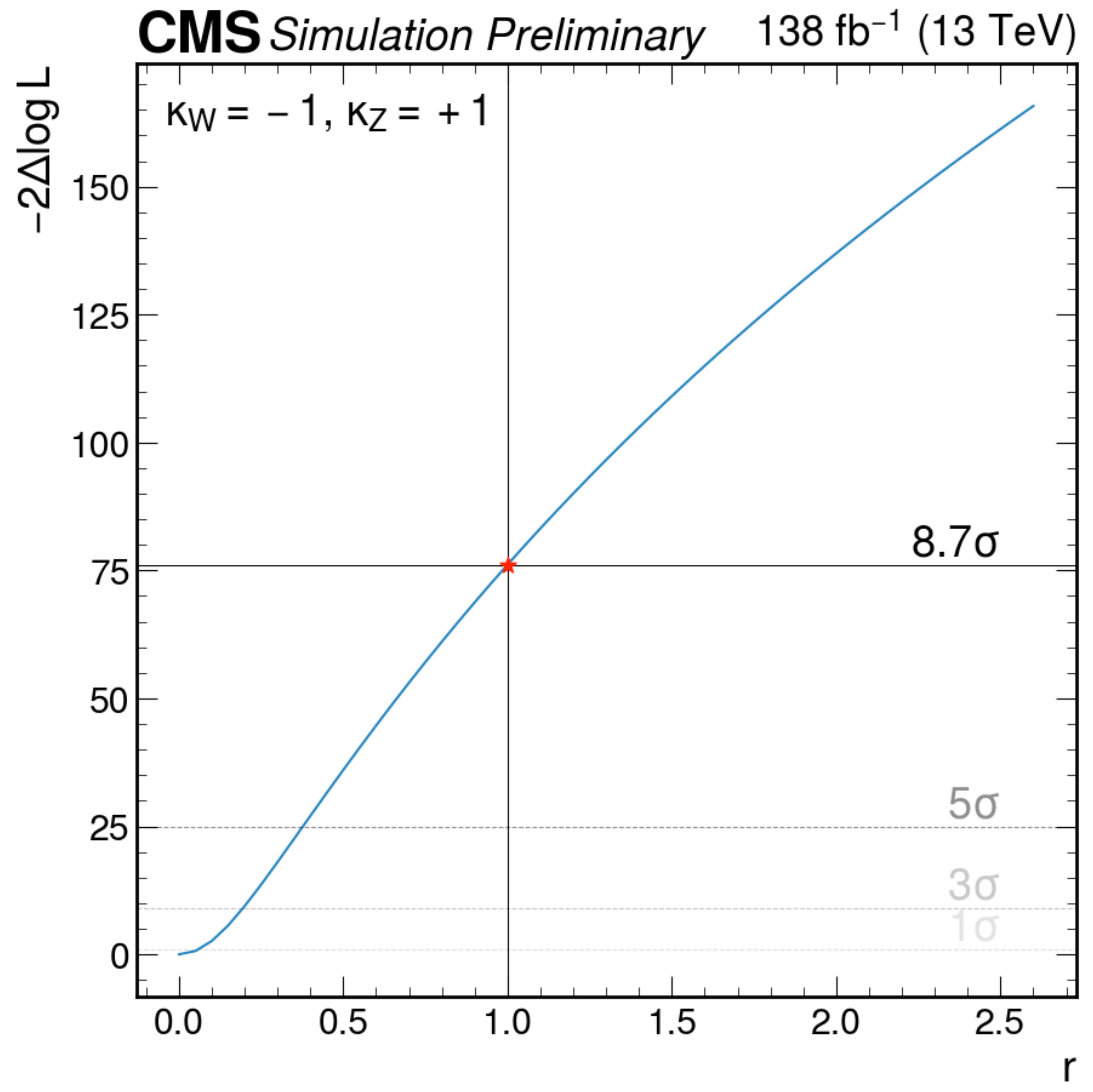
Region A | **Expected signal (λ_{WZ} = -1): 366±2.9**

Predicted background: 120±16.1±15.3

stat. *syst.*

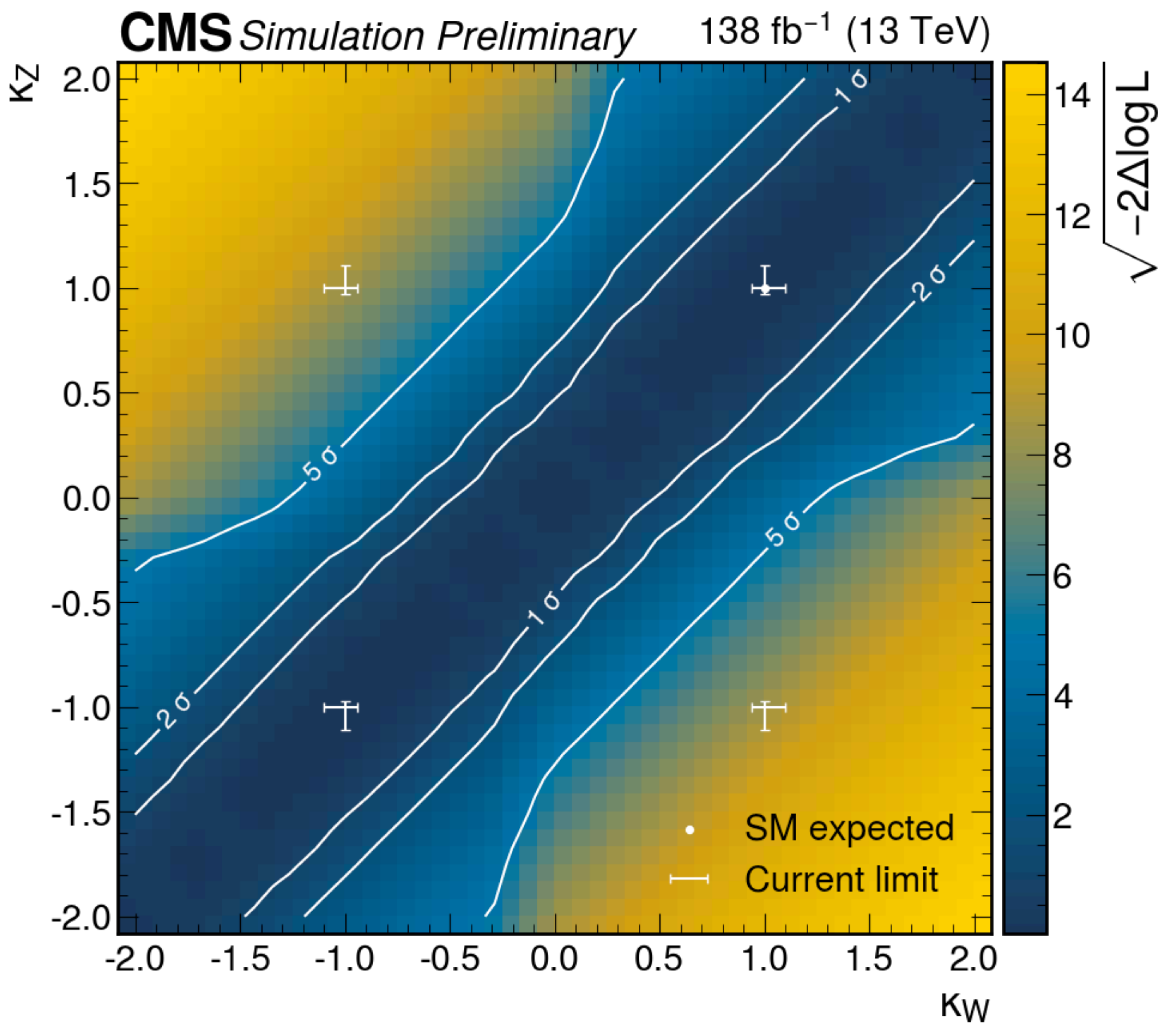
UC San Diego

Expected Results



Madgraph reweighting

Interpolated exclusion σ for each κ_W, κ_Z point



Expected signal ($\lambda_{WZ} = -1$): 366 ± 2.9

Predicted background: 120 ± 16.1 ± 15.3

stat. *syst.*

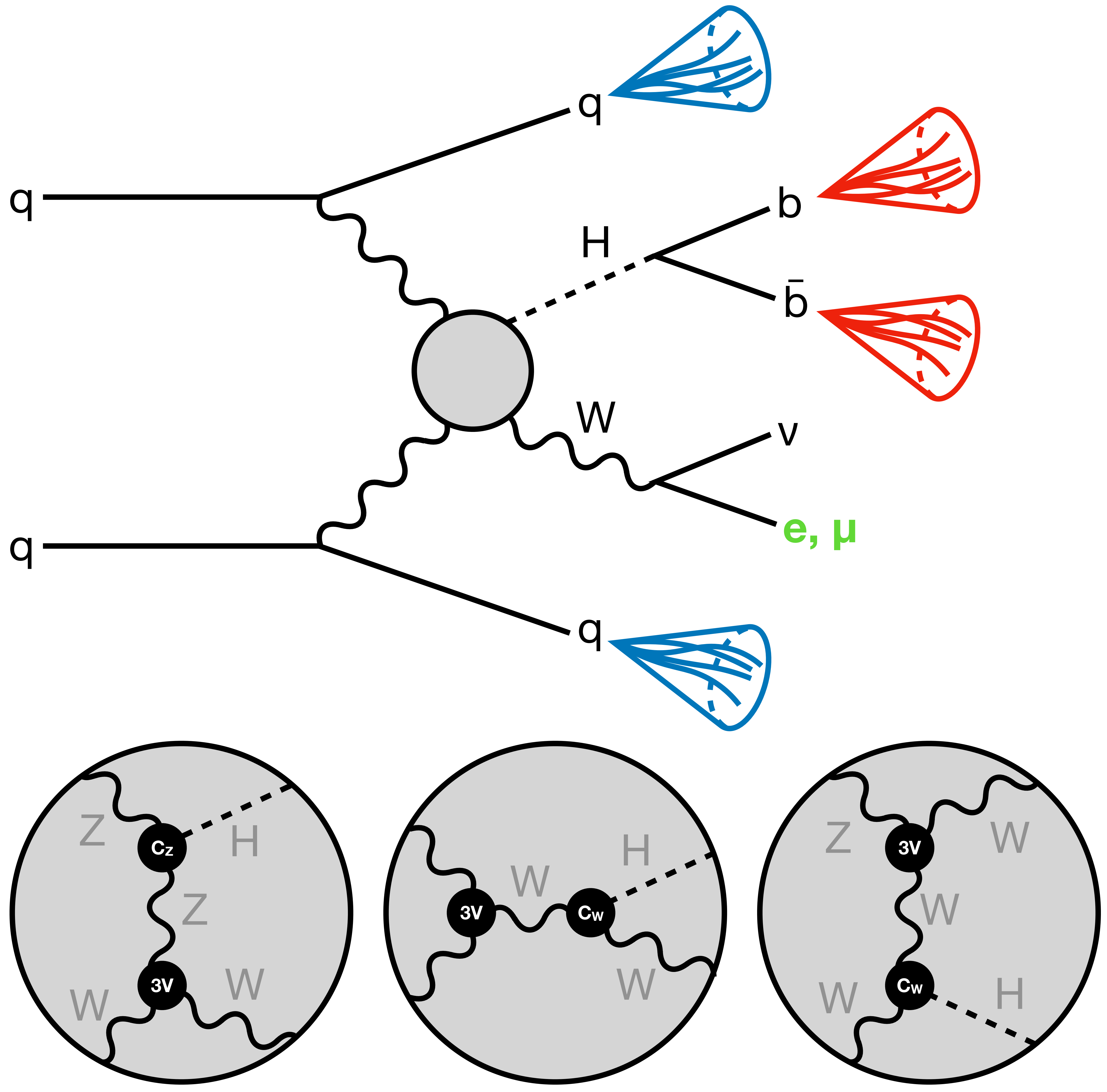
Strong exclusion of $\lambda_{WZ} < 0$ scenarios allowed by current limits

UC San Diego

Resolved Analysis

Analysis Strategy

1. Leverage High Level Triggers (HLTs) for most basic selection
   - We use the **single lepton** triggers
2. Construct a Signal Region (SR) for final fit
   - Train a BDT here to more efficiently select signal events
3. Construct a Control Region (CR) for MC validation and final fit
4. Perform two binned likelihood fits



(a) (b) (c)

UC San Diego

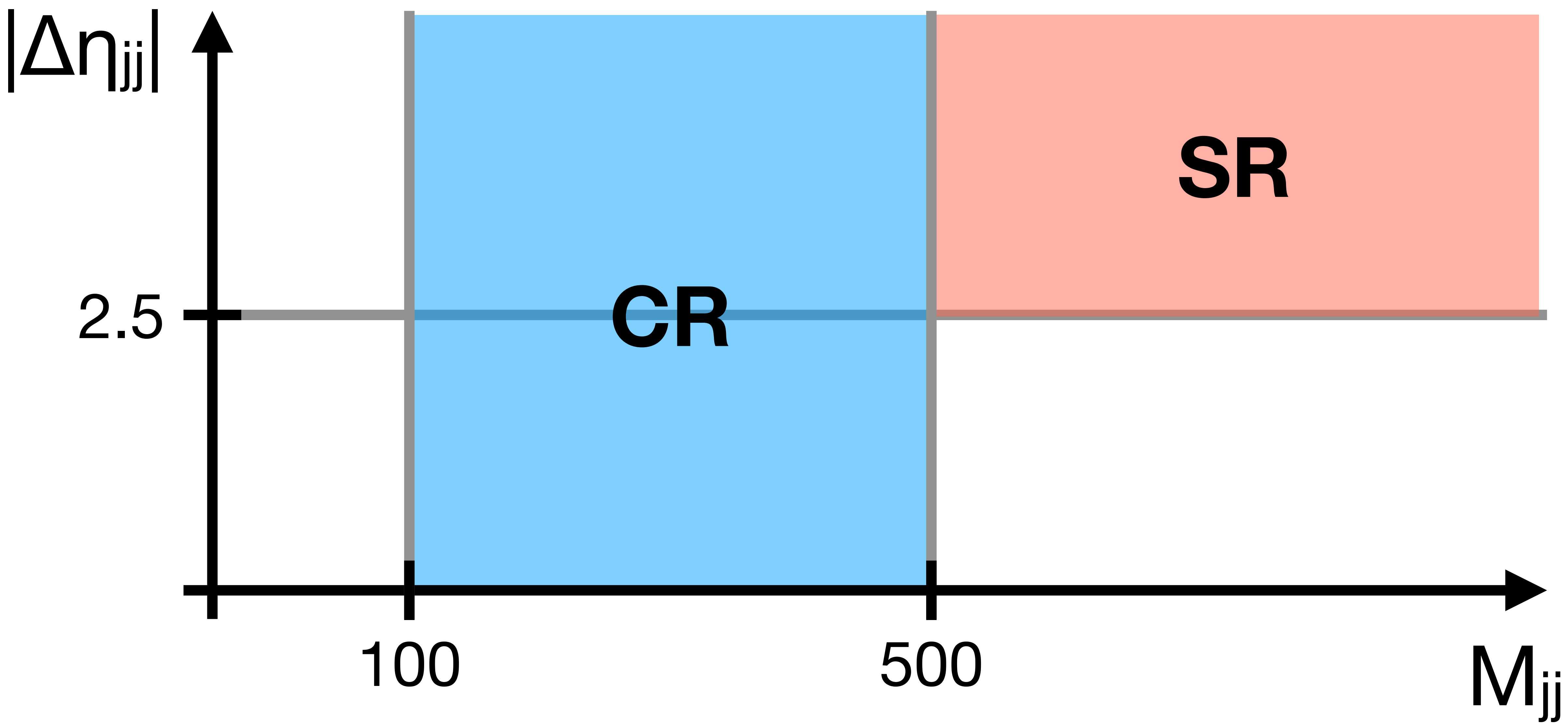
Signal and Control Regions

- Preselection applied to both CR and SR
- SR defined in large M_{jj} and $|\Delta\eta_{jj}|$ region
  - BDT trained here
- CR defined in low M_{jj} sideband
  - Data/MC agreement validated for bkg.
- **Both of these regions are used in the final fit**

Preselection:

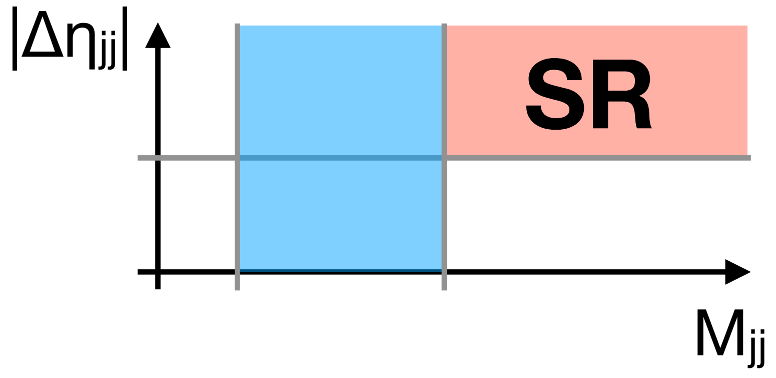
- Single lepton HLTs
- Basic object selections (VBS, $H \rightarrow b\bar{b}$ jets, 1 lepton)
- $p_T(W) = p_T(\ell)$ + MET > 35 GeV
- $M_{bb} \in [50, 150]$ GeV

CR vs. SR topography:



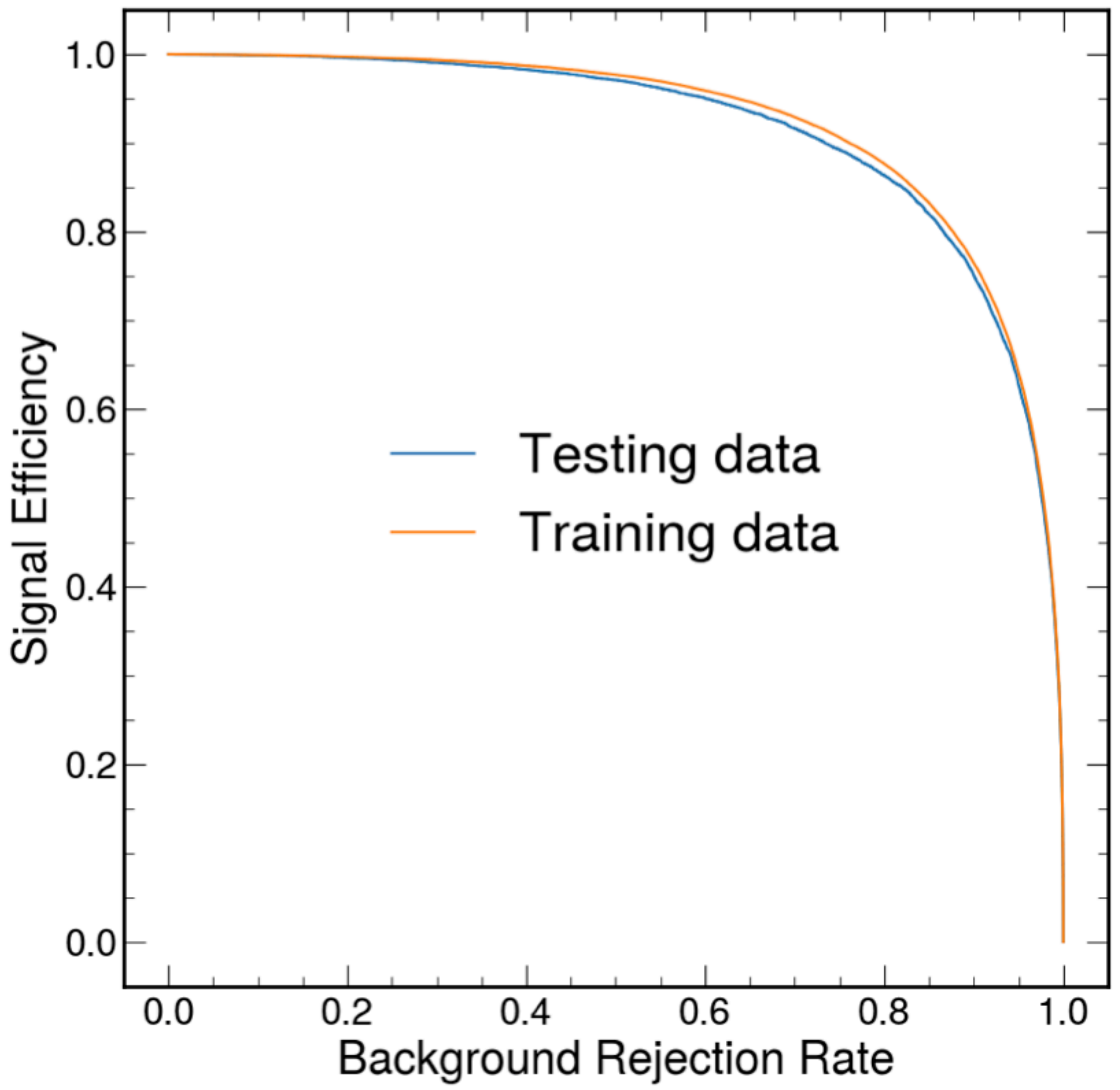
UC San Diego

SR: BDT Training

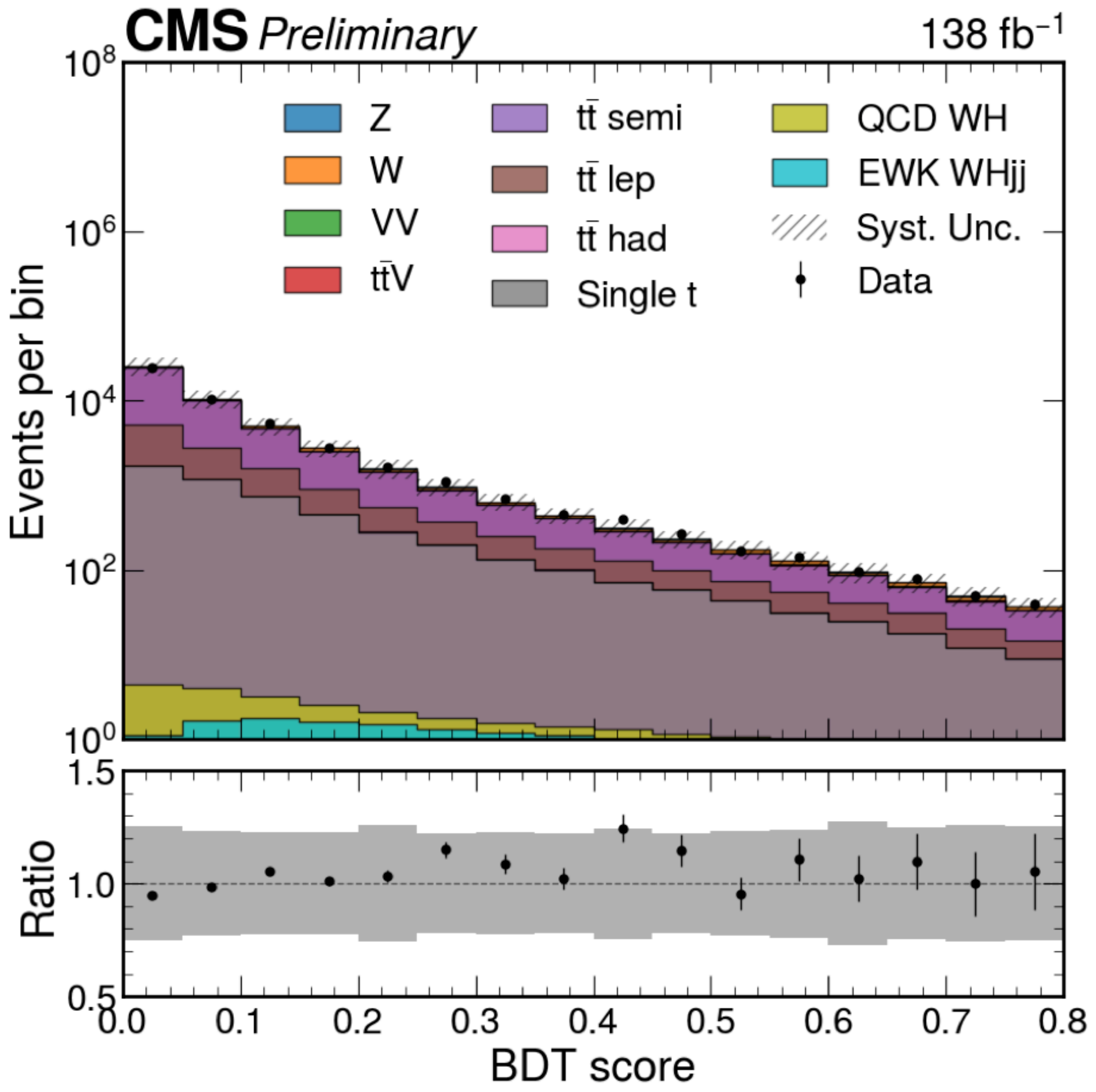


Input features

- The p_T, η, and ϕ of all selected leptons and jets
- The lepton charge, converted to a 0 or 1, where 0 corresponds to negative charge and 1 corresponds to positive charge
- The lepton flavor, converted to a 0 or 1, where 0 corresponds to muons and 1 corresponds to electrons
- The DeepJet b-tagging discriminator of all selected jets
- The MET p_T and ϕ
- The Higgs dijet mass
- The Higgs dijet $\Delta\eta$
- The VBF dijet mass
- The VBF dijet $\Delta\eta$
- The number of additional jets in the event
- The number of additional b-tagged jets in the event
- The transverse mass of the lepton plus jet systems for both b-tagged jets
- The ΔR between all pairs of jets



Good performance ✓

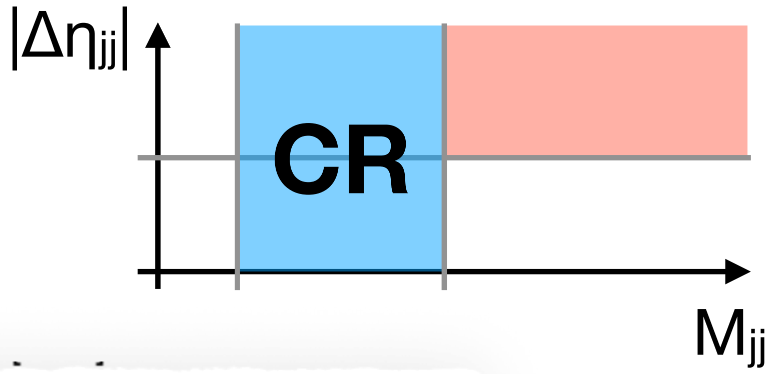


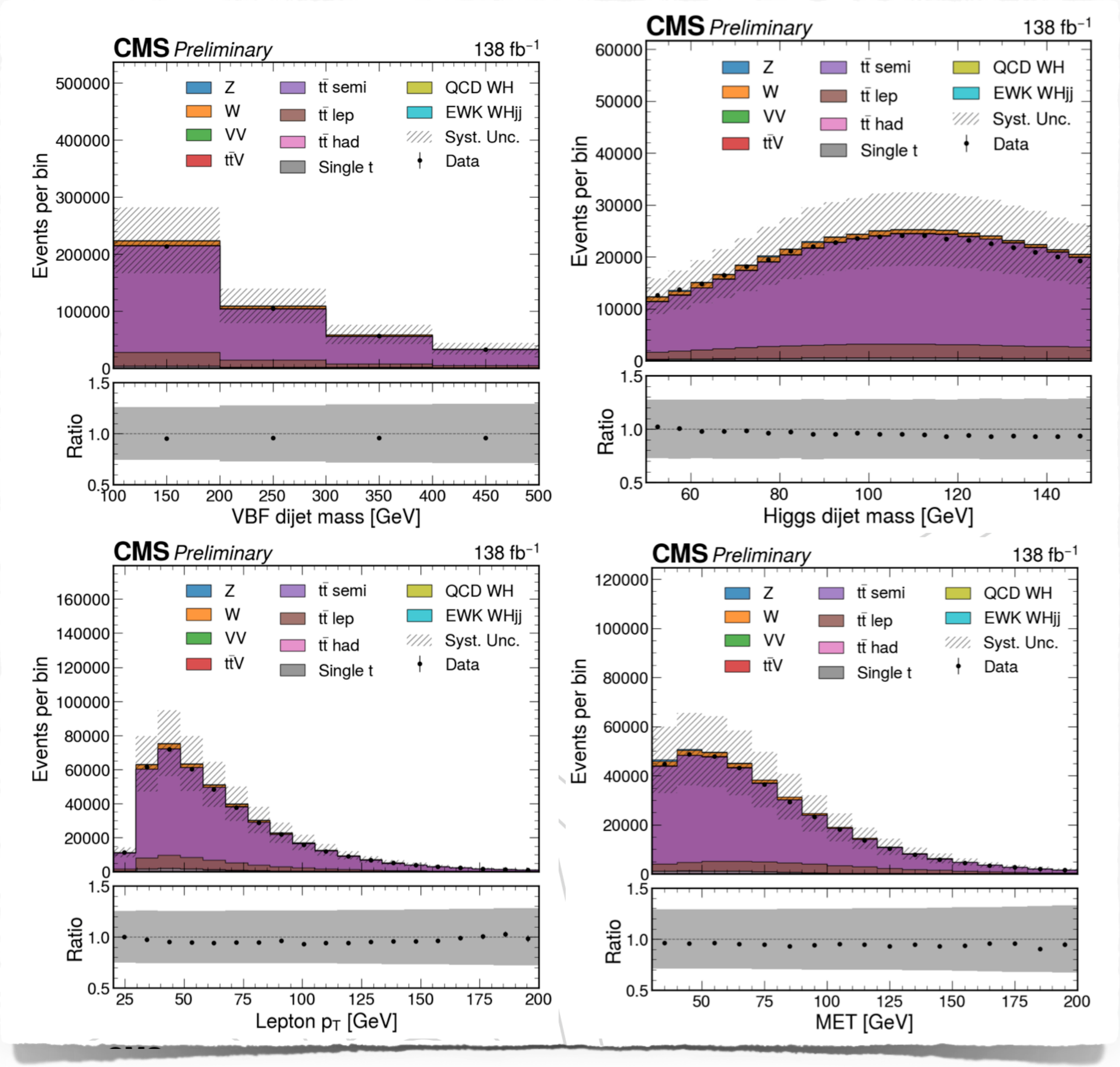
Good agreement ✓

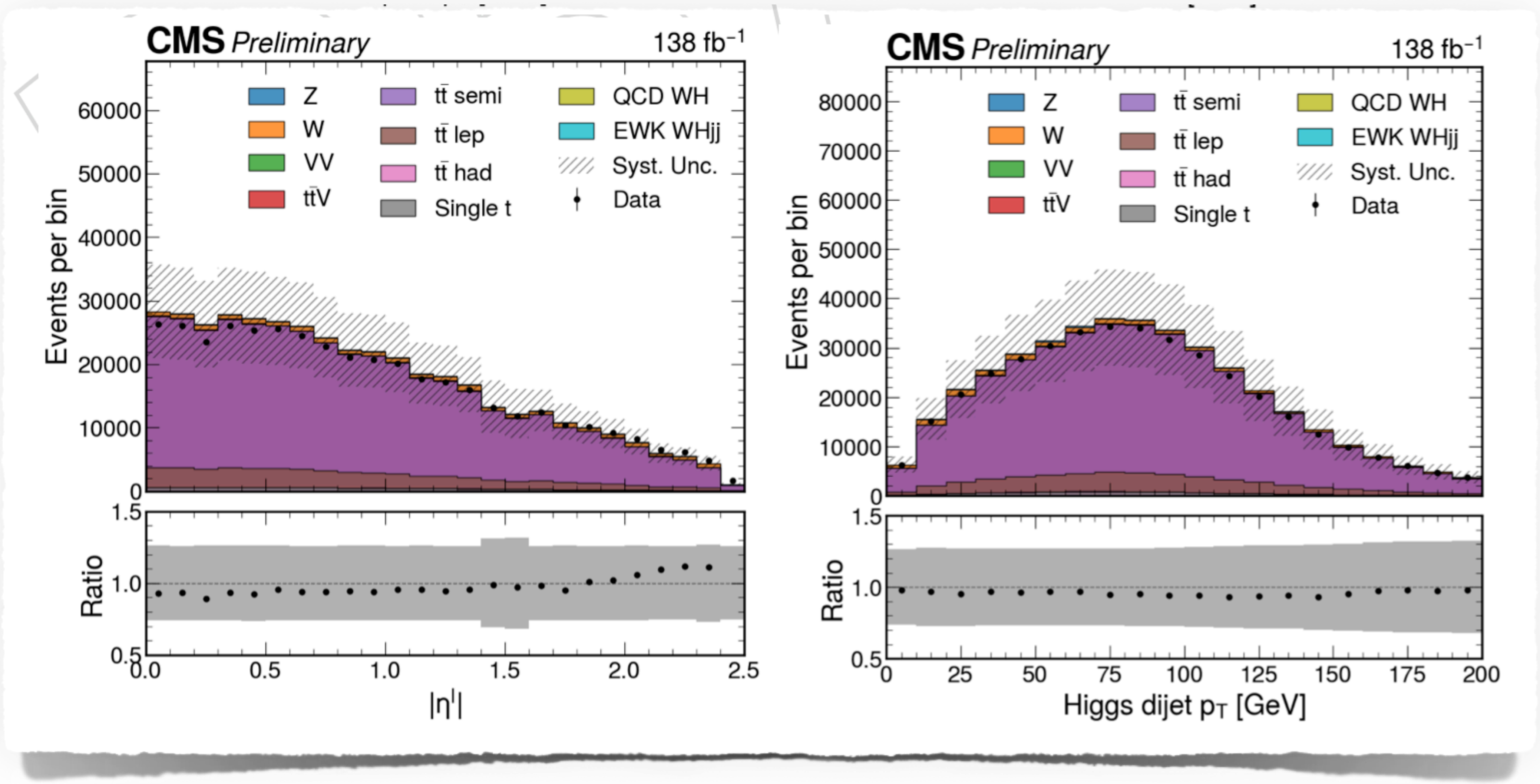
- Aforementioned signal features (plus a few more) used as input variables
  - Most important features: M_{bb}, M_{jj}, $N_{\text{extra jets}}$, $p_{T,\,b}$ (full feature ranking in backup)
- **Low BDT score (< 0.8) distribution validated against data**

UC San Diego

CR: MC Validation







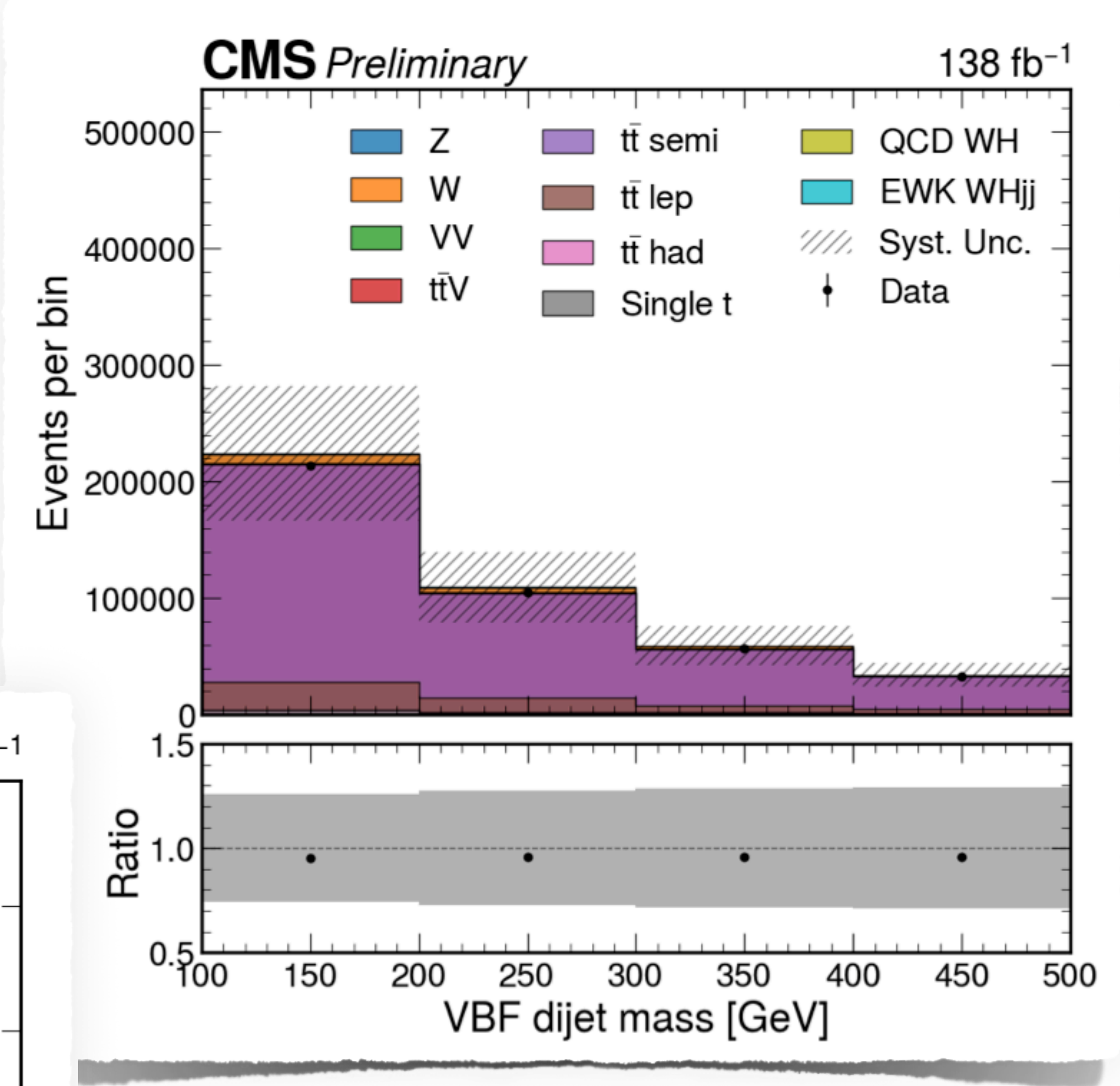
Good agreement ✓

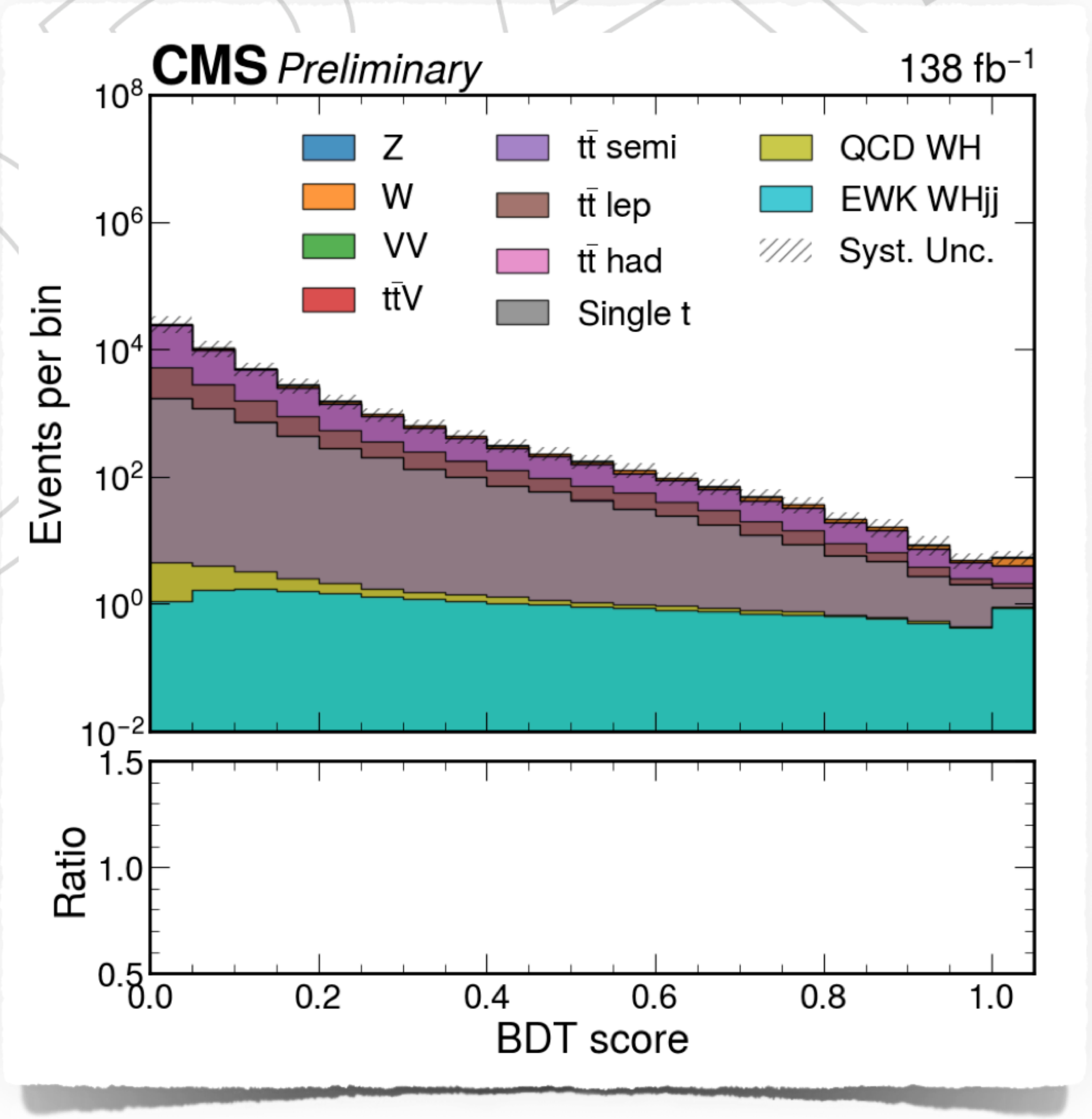
Take expected background from MC

UC San Diego

Expected Results

- Perform a binned likelihood fit on two distributions:
  - CR divided into 4 M_{jj} bins (100 GeV wide)
  - SR divided into 20 BDT score bins (0.05 wide)
- Expected significance: **0.49σ**





UC San Diego

Backup (boosted)

BSM Signal Models

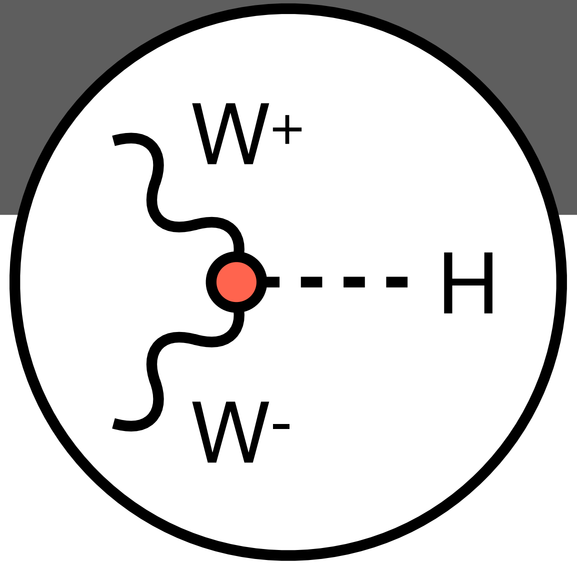
$\kappa_W = -1$

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})
```



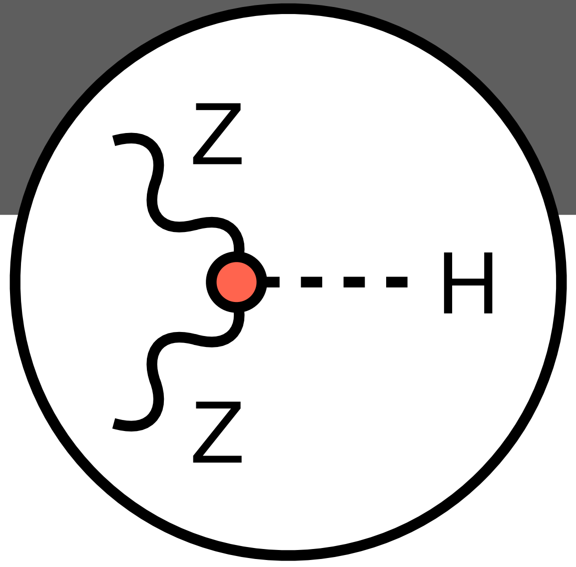
$\kappa_Z = -1$

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!

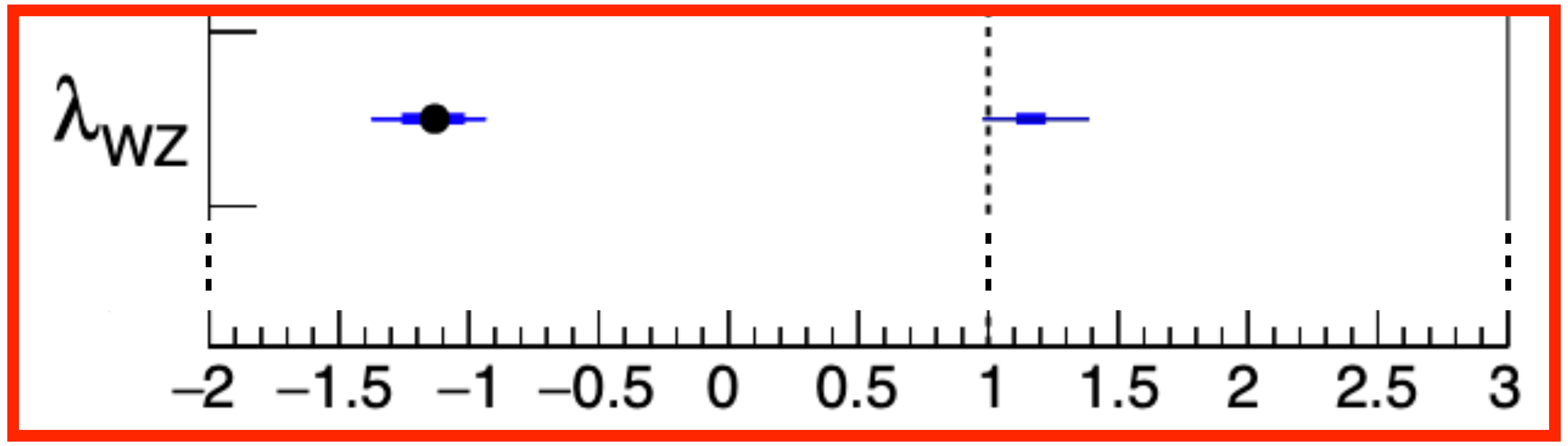
VBS WH Cross Sections

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

2016 CMS Result



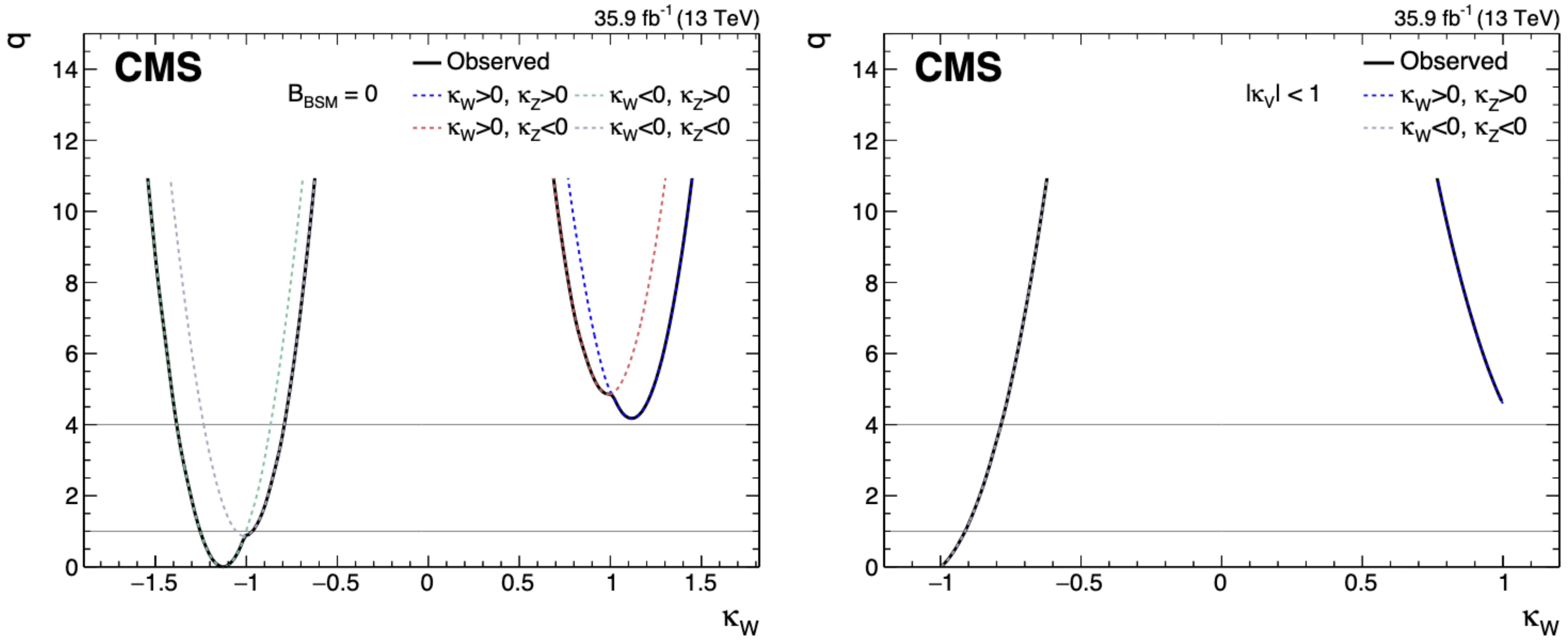


Figure 13: Scan of the test statistic q as a function of κ_W in the generic κ model assuming $\mathcal{B}_{BSM} = 0$ (left) and allowing $\mathcal{B}_{inv}$ and $\mathcal{B}_{undet}$ to float (right). The different colored lines indicate the value of q for different combinations of signs for κ_W and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

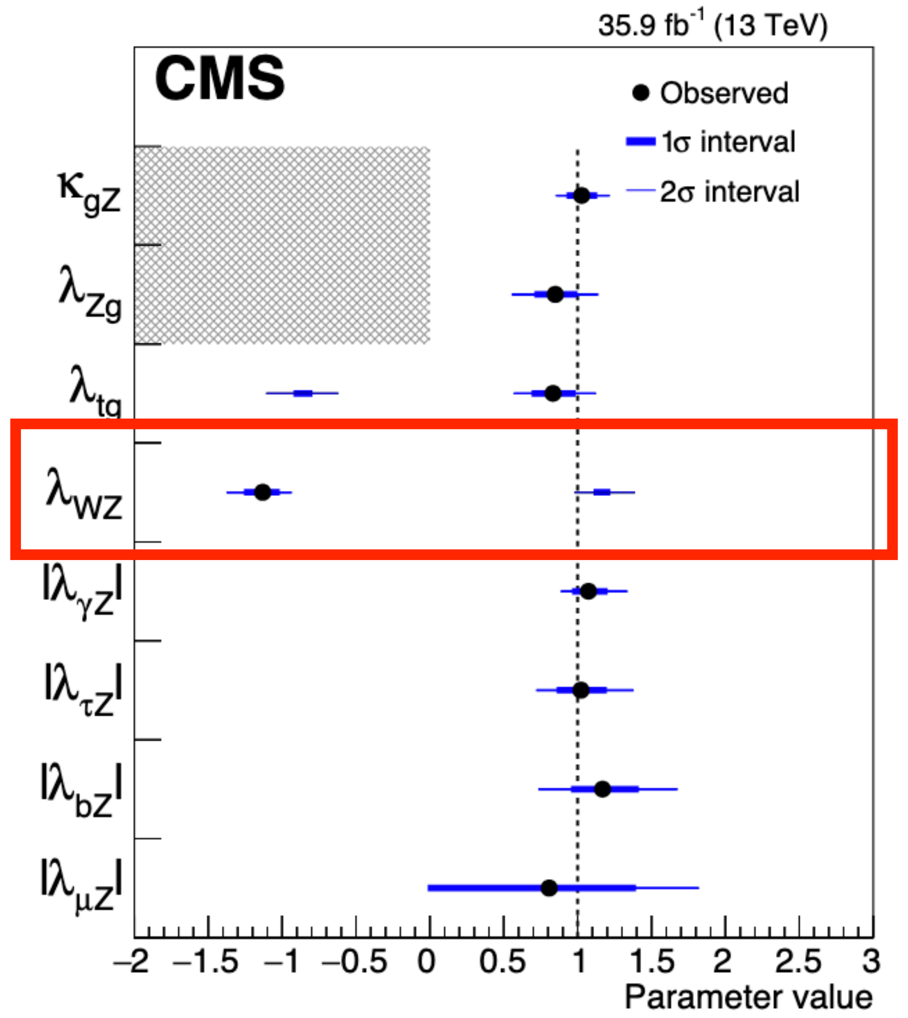
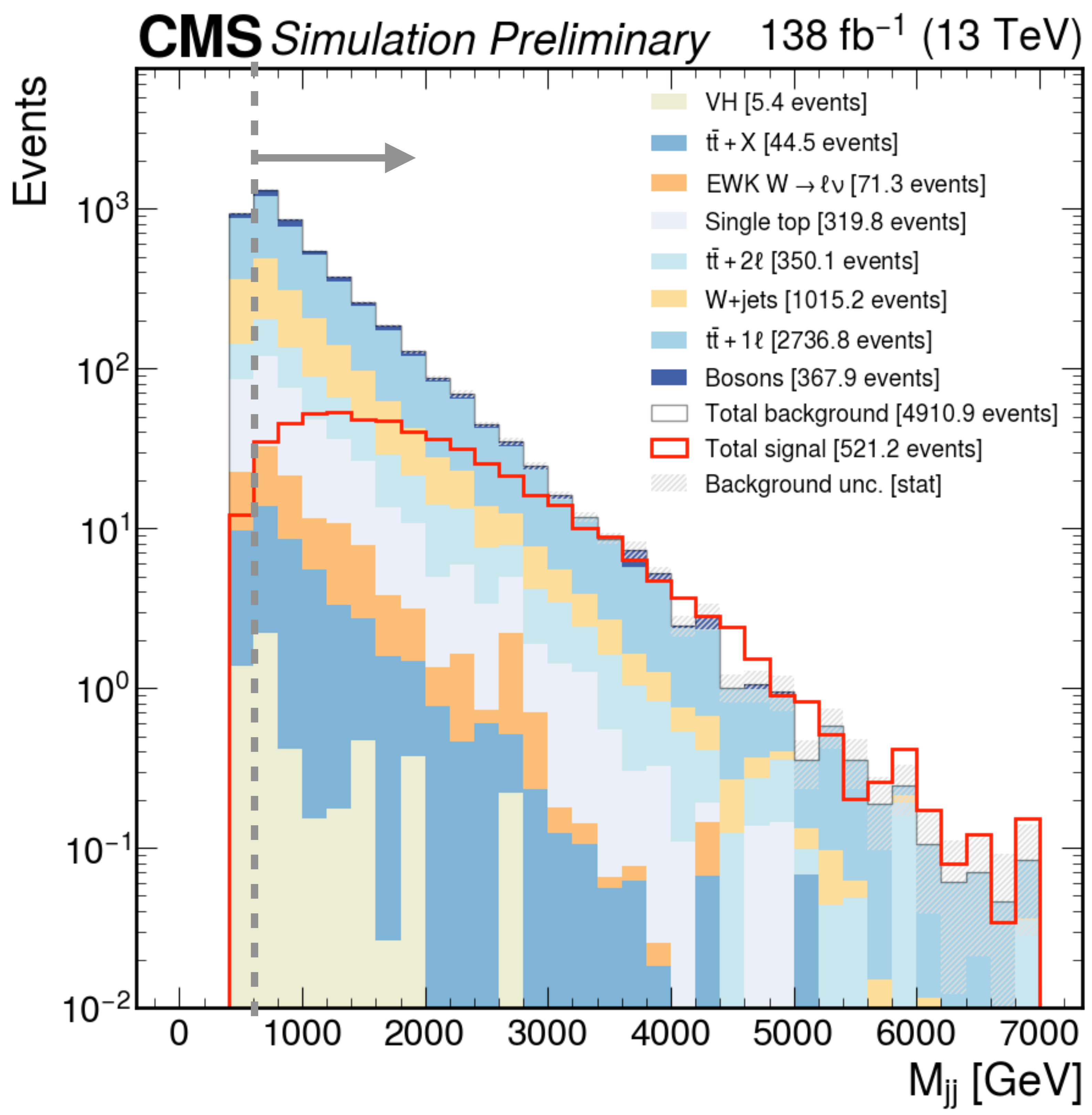
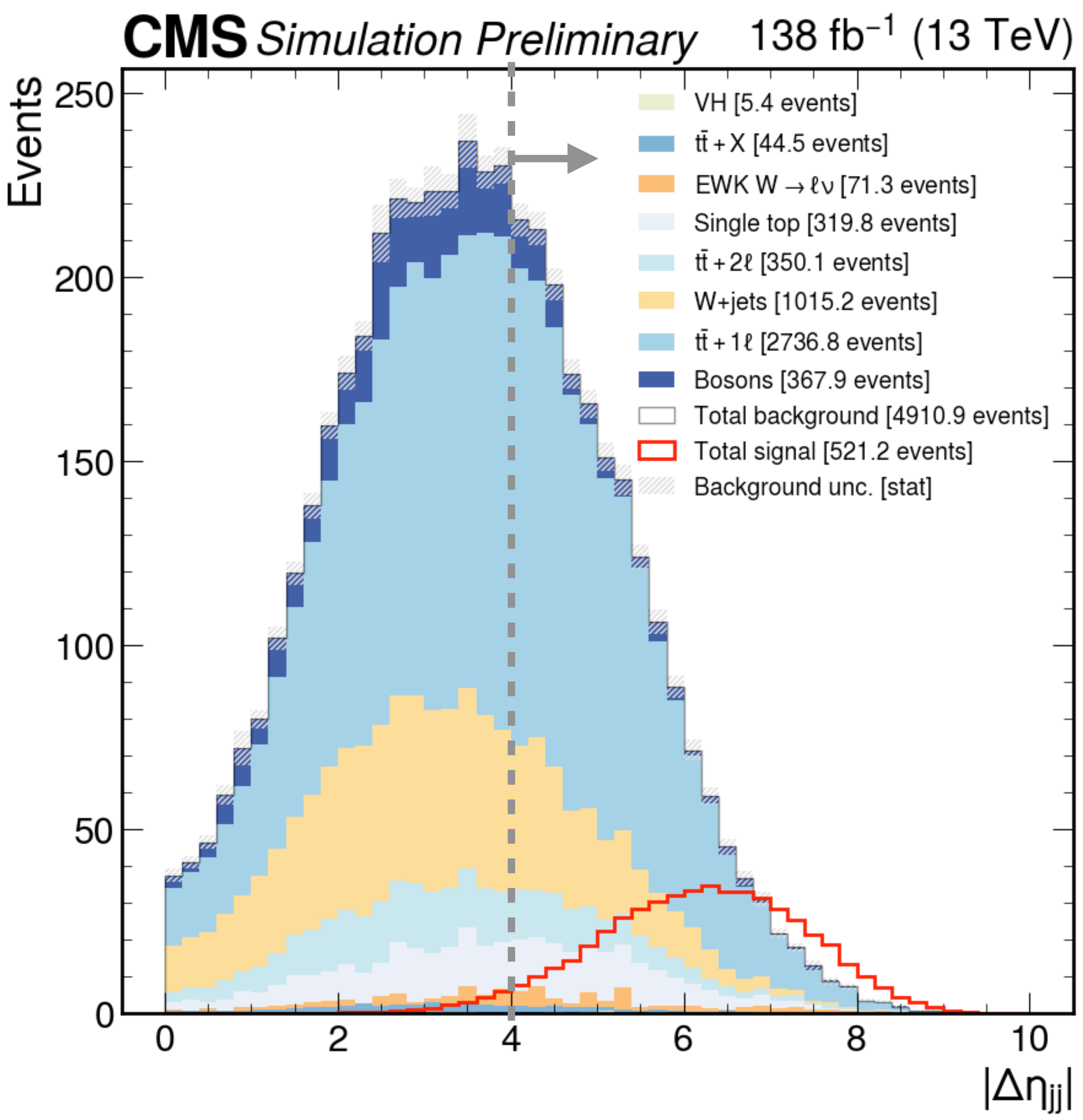


Figure 16: Summary of the model with coupling ratios and effective couplings for the ggH and $H \to \gamma\gamma$ loops. The points indicate the best fit values while the thick and thin horizontal bars show the 1σ and 2σ CL intervals, respectively. For this model, both positive and negative values of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

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

VBS Jets

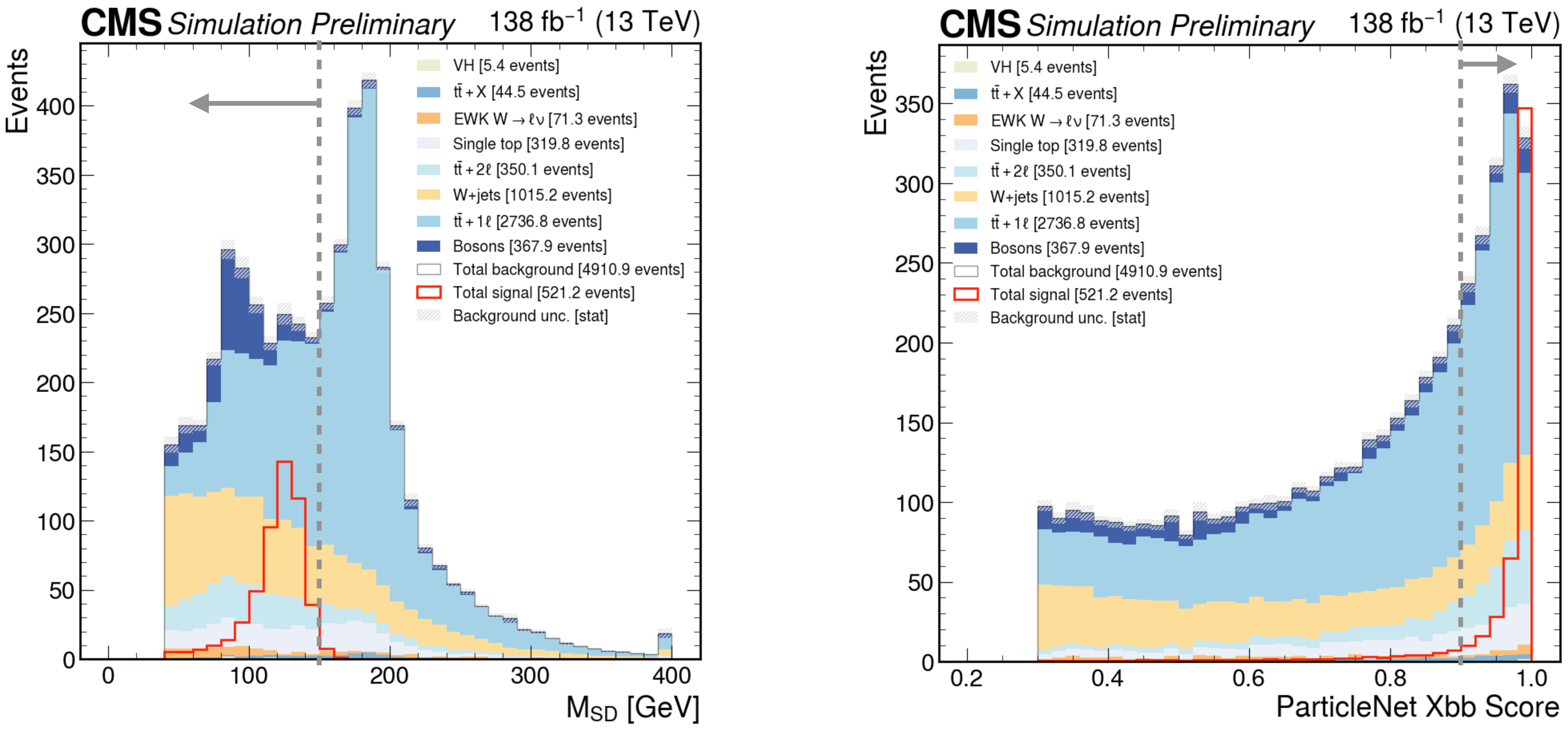




- M_{jj} = invariant mass of VBS system, $|\Delta\eta_{jj}| = \eta_1 - \eta_2$
- **VBS signature for signal** is clear: large M_{jj} & $|\Delta\eta_{jj}|$

UC San Diego

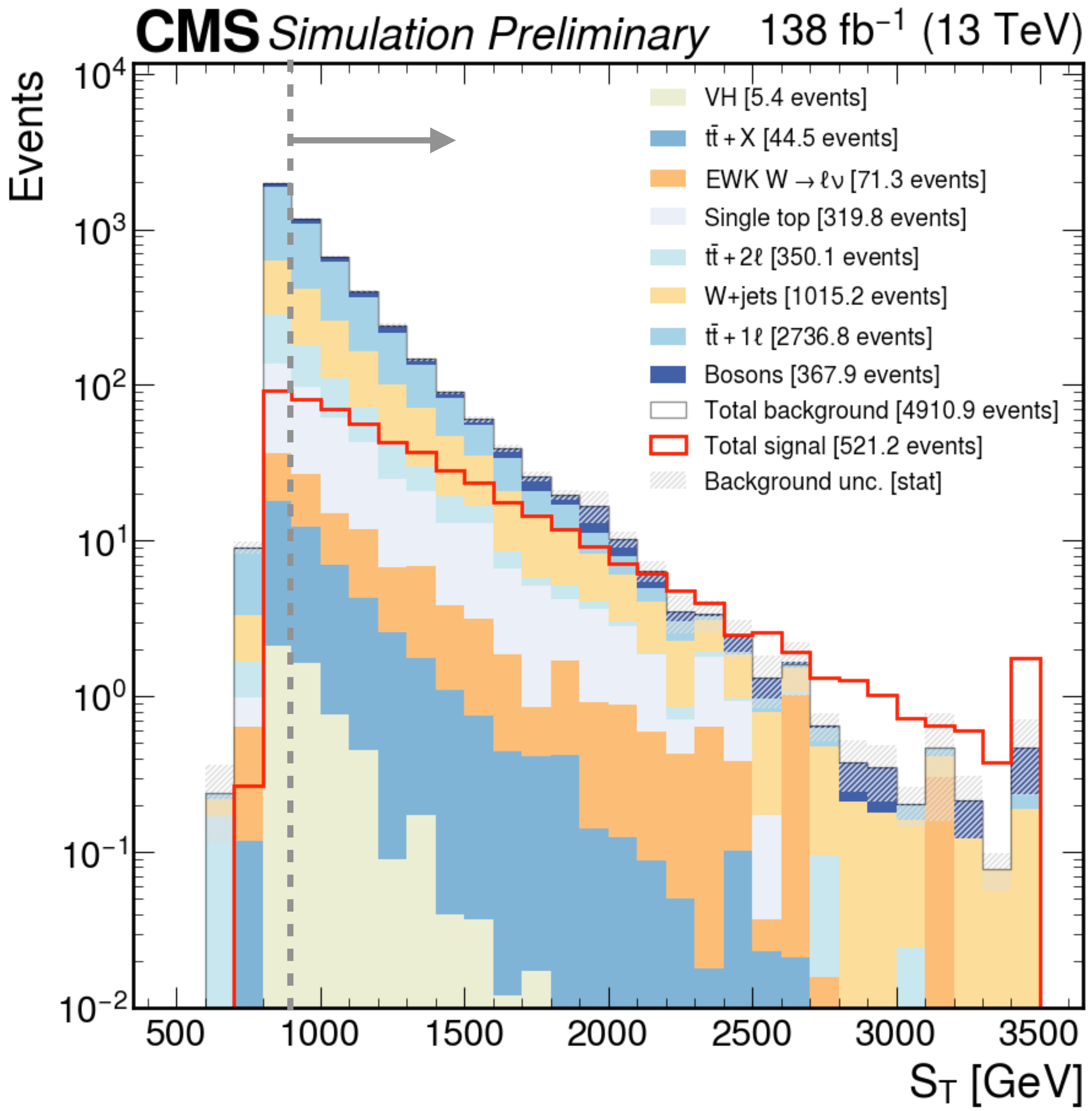
H→bb Large-radius Jet



- M_{SD} = "softdrop" mass of jet, ParticleNet Xbb = mass-decorrelated X→bb jet tagger
- Higgs peak + performant tagger gives **strong signal separation**

Boosted WH

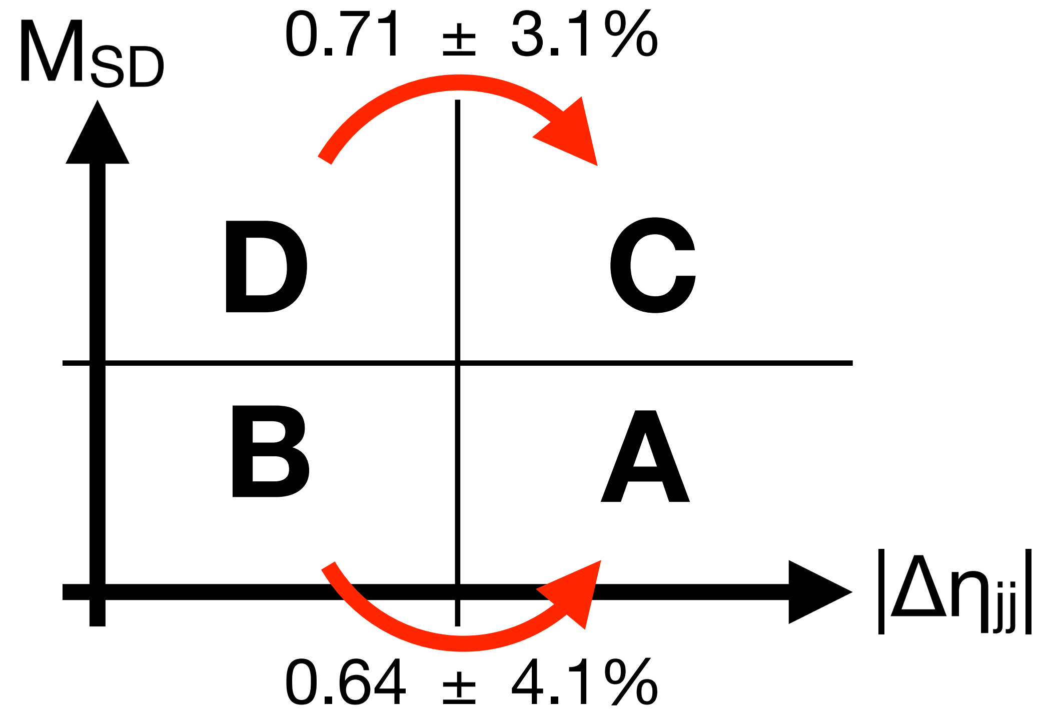
- $S_T = p_T(\ell) + MET + p_T(H \to b\bar{b})$
  - i.e. transverse energy of W + H
- Captures **boosted WH from λ_{WZ} = -1**
  - Large number of signal events in S_T tail
  - Background falls exponentially
- Signal is similarly boosted for most κ_W, κ_Z points where $\lambda_{WZ} < 0$



ABCD Systematic Error

Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} ≥ 150 GeV	D	172.97	3.25	6.92	0.40	142	11.92
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} ≥ 150 GeV	C	241.93	5.83	0.27	0.08	201	14.18
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} < 150 GeV	B	181.10	4.40	11.62	0.52	170	13.04
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} < 150 GeV (SR)	A	116.41	3.84	366.30	2.92	–	–



- Errors: 10% (syst.), 13% (stat.)

Over-predicted

Bkg comp. syst. (backup)

$$D^{pred}_{MC} = \frac{A_{MC}}{B_{MC}} \times C_{MC} = \mathbf{129.48}$$

$$\epsilon_{syst} = \left|1 - \frac{D^{pred}_{MC}}{D_{MC}}\right| = \left|1 - \frac{129.5}{116.4}\right| = 11\% \oplus 6\% = 13\%$$

$$D^{pred}_{data} = \frac{A_{data}}{B_{data}} \times C_{data} = 120.10$$

$$\epsilon_{stat} = \sqrt{\left(\frac{\sqrt{A_{data}}}{A_{data}}\right)^2 + \left(\frac{\sqrt{B_{data}}}{B_{data}}\right)^2 + \left(\frac{\sqrt{C_{data}}}{C_{data}}\right)^2} = \sqrt{\frac{1}{A_{data}} + \frac{1}{B_{data}} + \frac{1}{C_{data}}} = 13\%$$

Predicted SR Yield: 120.1 ± 16.07 ± 15.30

stat. *syst.*

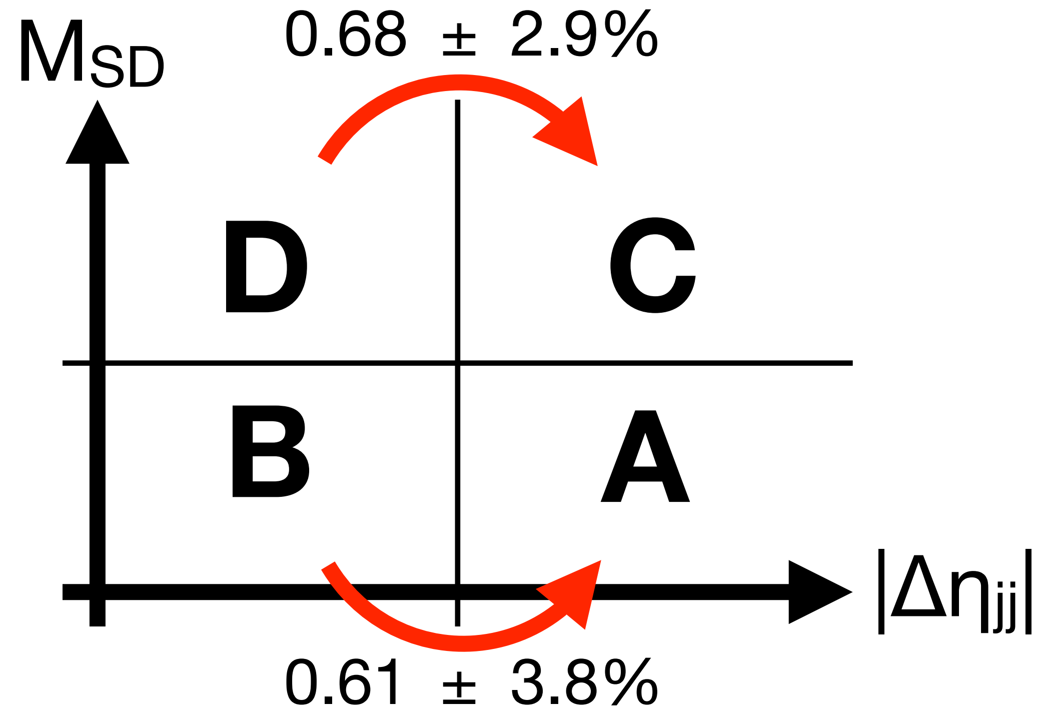
*err = $\sqrt{(\Sigma_i w_i^2)}$ for MC, $\sqrt{(\text{count})}$ for data

UC San Diego

ABCD W+jets Composition

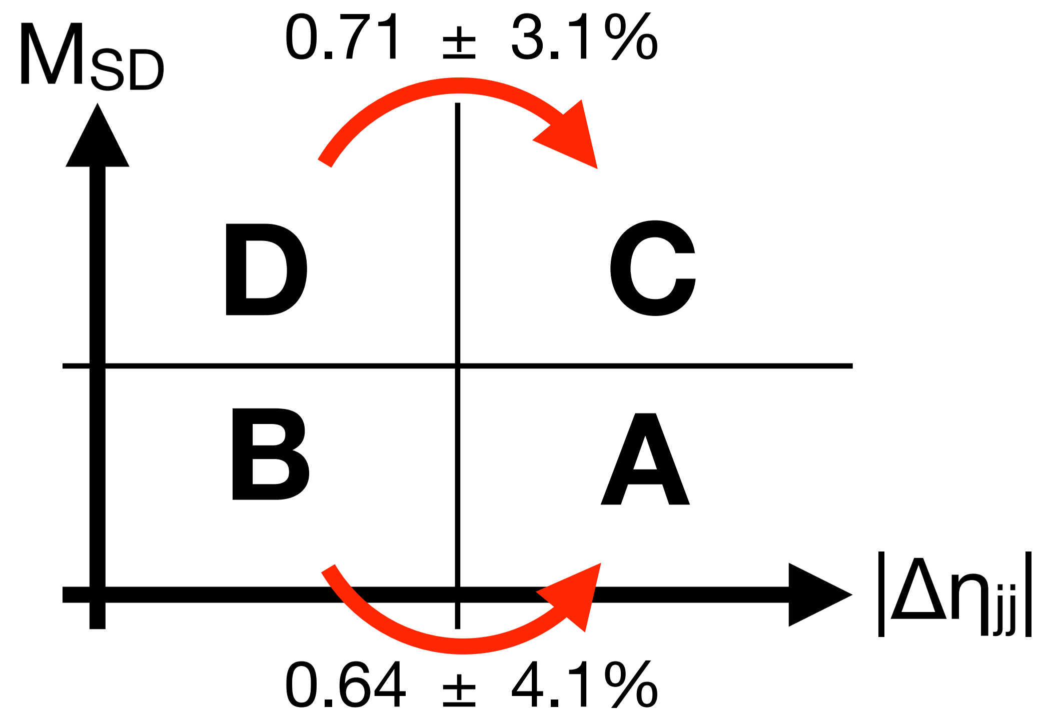
Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9 (WJets x 2)

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} \geq 150$ GeV	D	184.26	3.48	6.92	0.40	142	11.92
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} \geq 150$ GeV	C	272.50	5.98	0.27	0.08	201	14.18
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} < 150$ GeV	B	223.95	4.72	11.62	0.52	170	13.04
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} < 150$ GeV (SR)	A	137.64	4.42	366.30	2.92	—	—



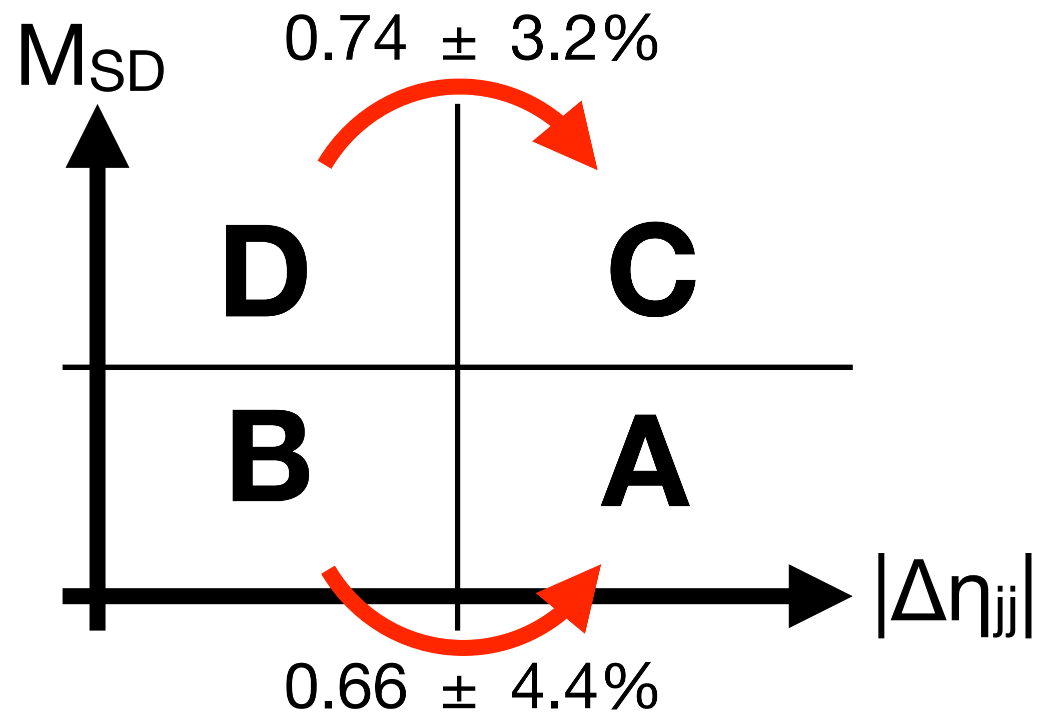
Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} \geq 150$ GeV	D	172.97	3.25	6.92	0.40	142	11.92
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} \geq 150$ GeV	C	241.93	5.83	0.27	0.08	201	14.18
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} < 150$ GeV	B	181.10	4.40	11.62	0.52	170	13.04
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} < 150$ GeV (SR)	A	116.41	3.84	366.30	2.92	—	—



Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9 (WJets x 0.5)

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} \geq 150$ GeV	D	167.32	3.19	6.92	0.40	142	11.92
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} \geq 150$ GeV	C	226.65	5.79	0.27	0.08	201	14.18
$\lvert\Delta\eta_{jj}\rvert \leq 4$ AND $M_{SD} < 150$ GeV	B	159.67	4.32	11.62	0.52	170	13.04
$\lvert\Delta\eta_{jj}\rvert > 4$ AND $M_{SD} < 150$ GeV (SR)	A	105.79	3.68	366.30	2.92	—	—



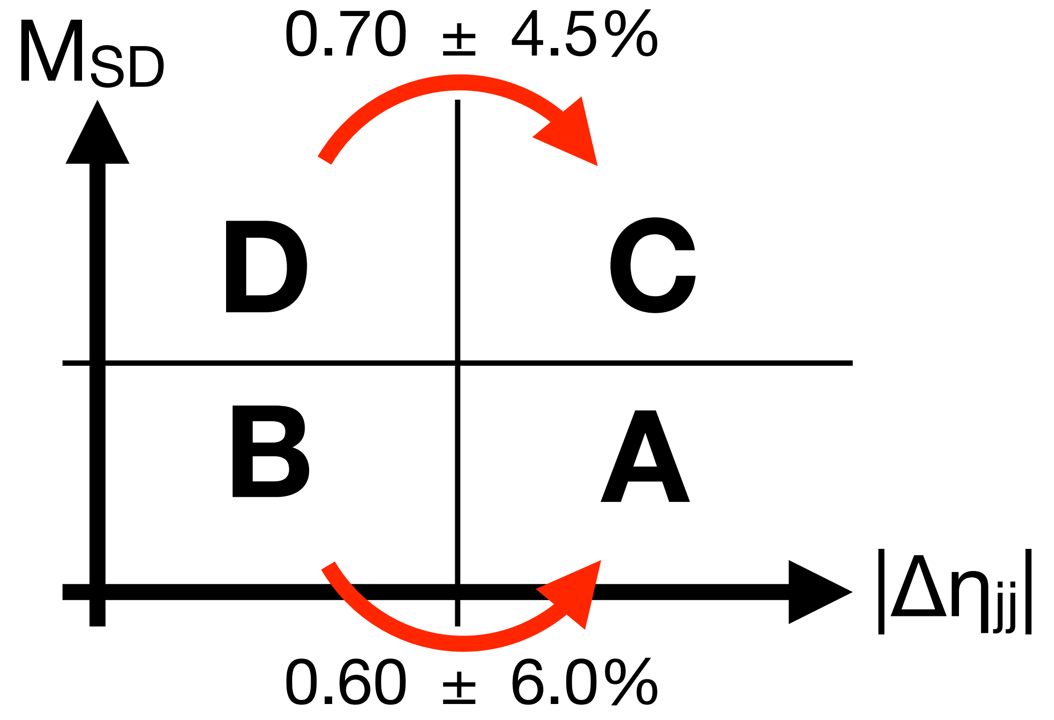
5.4% systematic

*err = $\sqrt{(\Sigma_i w_i^2)}$ for MC, $\sqrt{(\text{count})}$ for data

ABCD Bosons Composition

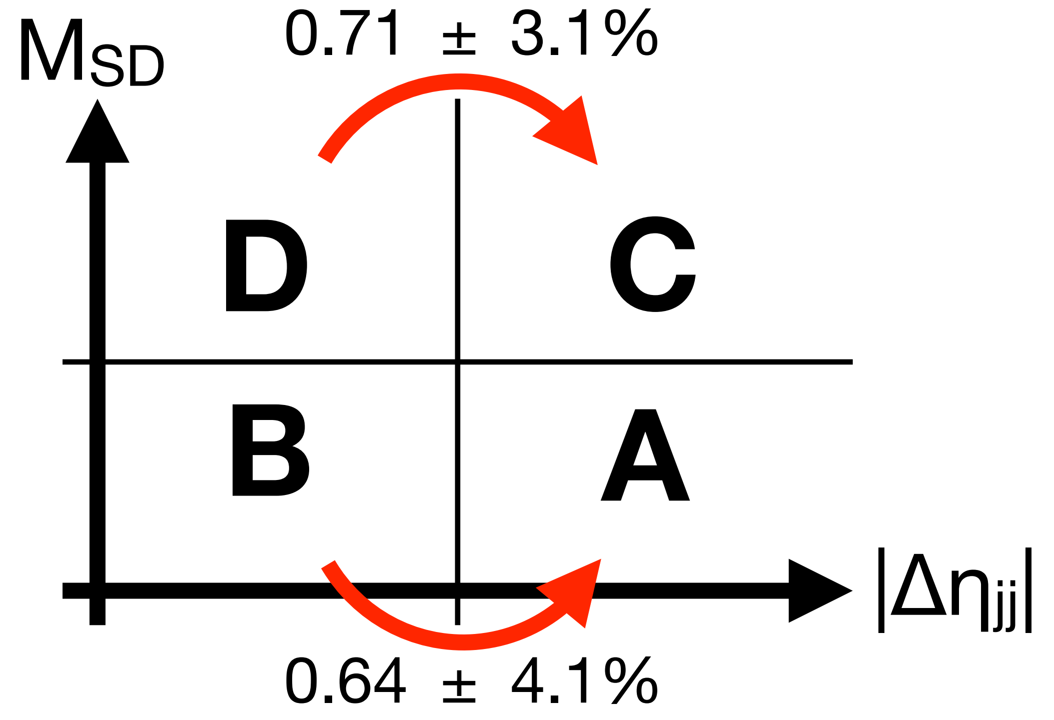
Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9 (Bosons x 2)

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} ≥ 150 GeV	D	173.96	3.46	6.92	0.40	142	11.92
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} ≥ 150 GeV	C	249.87	9.99	0.27	0.08	201	14.18
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} < 150 GeV	B	202.63	7.17	11.62	0.52	170	13.04
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} < 150 GeV (SR)	A	122.39	5.97	366.30	2.92	—	—



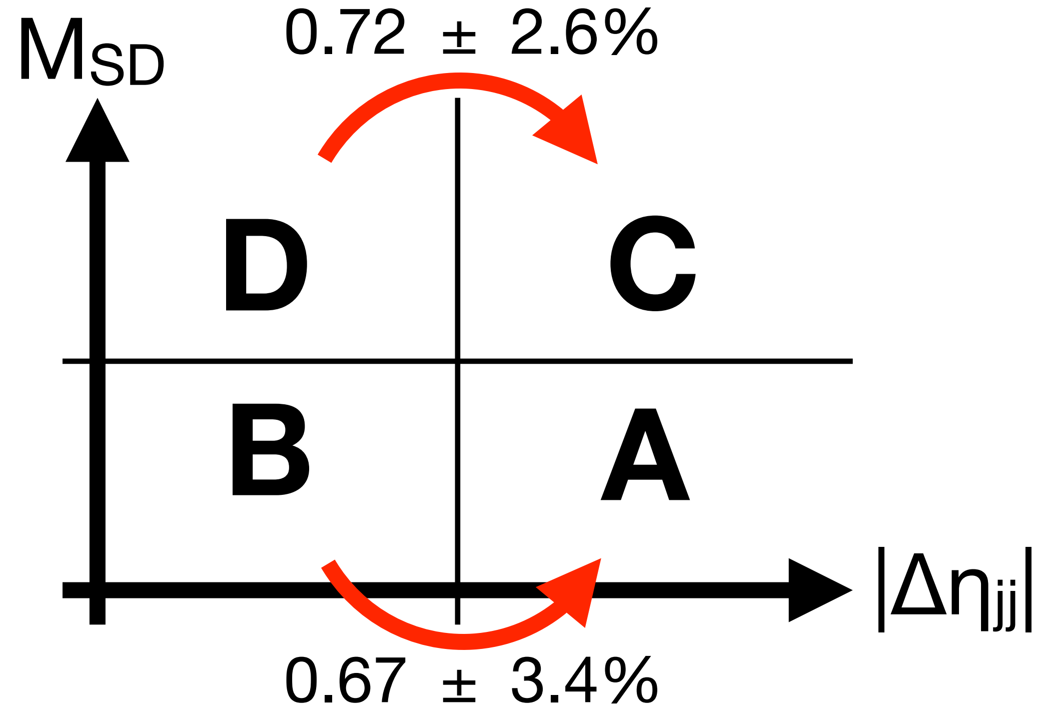
Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} ≥ 150 GeV	D	172.97	3.25	6.92	0.40	142	11.92
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} ≥ 150 GeV	C	241.93	5.83	0.27	0.08	201	14.18
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} < 150 GeV	B	181.10	4.40	11.62	0.52	170	13.04
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} < 150 GeV (SR)	A	116.41	3.84	366.30	2.92	—	—



Preselection AND M_{jj} > 600 GeV AND S_T > 900 GeV AND PNet Xbb > 0.9 (Bosons x 0.5)

Cut	Region	Bkg. (wgt)	Bkg. Err.*	Sig. (wgt)	Sig. Err.*	Data	Data Err.*
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} ≥ 150 GeV	D	172.47	3.20	6.92	0.40	142	11.92
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} ≥ 150 GeV	C	237.97	4.18	0.27	0.08	201	14.18
$\|\Delta\eta_{jj}\|$ ≤ 4 AND M_{SD} < 150 GeV	B	170.33	3.38	11.62	0.52	170	13.04
$\|\Delta\eta_{jj}\|$ > 4 AND M_{SD} < 150 GeV (SR)	A	113.42	3.08	366.30	2.92	—	—



*err = $\sqrt{(\Sigma_i w_i^2)}$ for MC, $\sqrt{(\text{count})}$ for data

2.6% systematic

Signal Systematics

- For most systematics:

1. Get nominal yield in SR
2. Get yield in SR after applying up/down variation
    - If scale factor: first divide each event weight by nominal value
3. Systematic = largest % difference in yield

Step 1

$$\text{yield} = y = \sum_{i=1}^{N} W_i$$

$$\text{where } W_i = \prod_j \omega_i$$

Step 2

$$y_{var} = \sum_{i=1}^{N} W_i \times \frac{\omega_{var}}{\omega_{nom}}$$

Step 3

$$\delta_{var} = \left| 1 - \frac{y_{var}}{y} \right|$$

$$\text{syst.} = \max\left(\delta_{up}, \delta_{down}\right)$$

Systematic	Size
PDF variations	2.2%
μ_F scale	17.5%
Parton shower ISR weights	0.6%
Parton shower FSR weights	1.7%
Pileup reweighting	0.2%
Pileup jet ID	0.8%
L1 pre-fire corrections	0.9%
Single-electron HLT scale factors	0.7%
Single-muon HLT scale factors	0.1%
Simulation stat. unc.	0.8%
Electron ID scale factors	1.4%
Muon ID scale factors	0.1%
Electron reco. scale factors	0.3%
Muon iso. scale factors	0.0%
ParticleNet Xbb scale factors	1.3%
DeepJet b-tagging scale factors	0.2%
MET unc.	0.1%
Jet energy scale	7.0%
Jet energy resolution	0.4%
Luminosity	1.6%
$H \to b\bar{b}$ BR	1.3%

Signal Systematics: PDF Variations

- For (Hessian) PDF systematics:
    1. Get overall variation/inclusive ratio for each of the 100 PDF variations
    2. Get nominal yield in SR
    3. Get yield in SR after applying a given variation
    4. Systematic = % difference in yield for each variation added in quadrature

Step 1

$$R_{var} = \frac{\sum_{i=0}^{N} \omega_i^{var}}{\sum_{i=0}^{N} \omega_i^{gen}}$$

Step 2

$$\text{yield} = y = \sum_{i=1}^{N} W_i$$

$$\text{where } W_i = \prod_j \omega_j$$

Step 3

$$y_{var} = \sum_{i=1}^{N} W_i \times \frac{\omega_i^{var}}{R_{var}}$$

Step 4

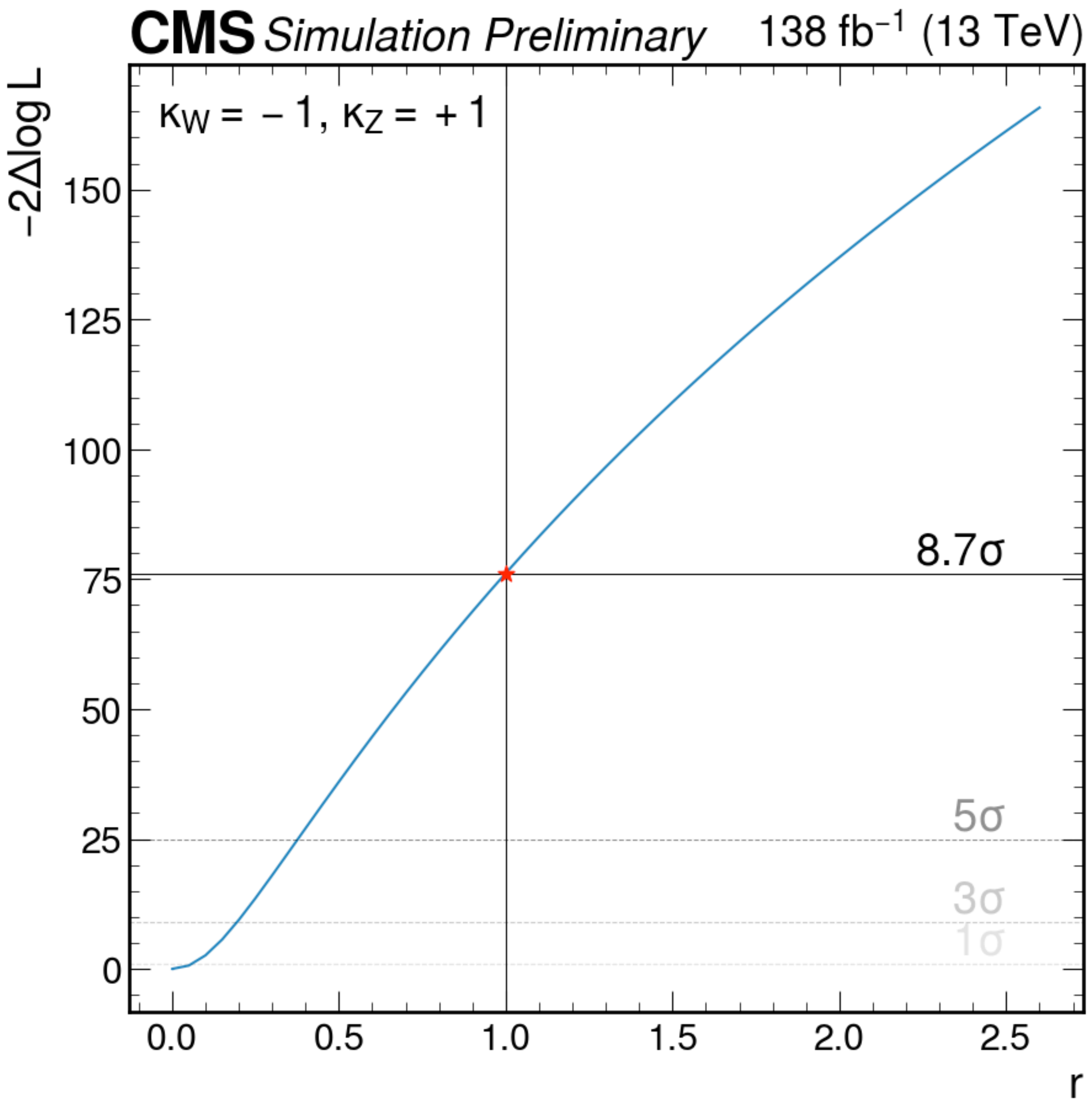
$$\delta_{var} = \left| 1 - \frac{y_{var}}{y} \right|$$

$$\text{syst.} = \left[\sum_{var} \delta_{var}^2 \right]^{1/2}$$

UC San Diego

Detailed Results

- Assessed rigorous set of systematics on signal simulation (backup)
- Maximum-likelihood fit for bkg-only hypothesis
  - “Observed” yield is artificially set to be equal to the predicted bkg. yield
  - Used CMS statistical tool*
- **We expect to exclude λ_{WZ} = -1 at 9σ**
- Waiting for internal approval to “unblind” analysis (i.e. look at the data in the SR)



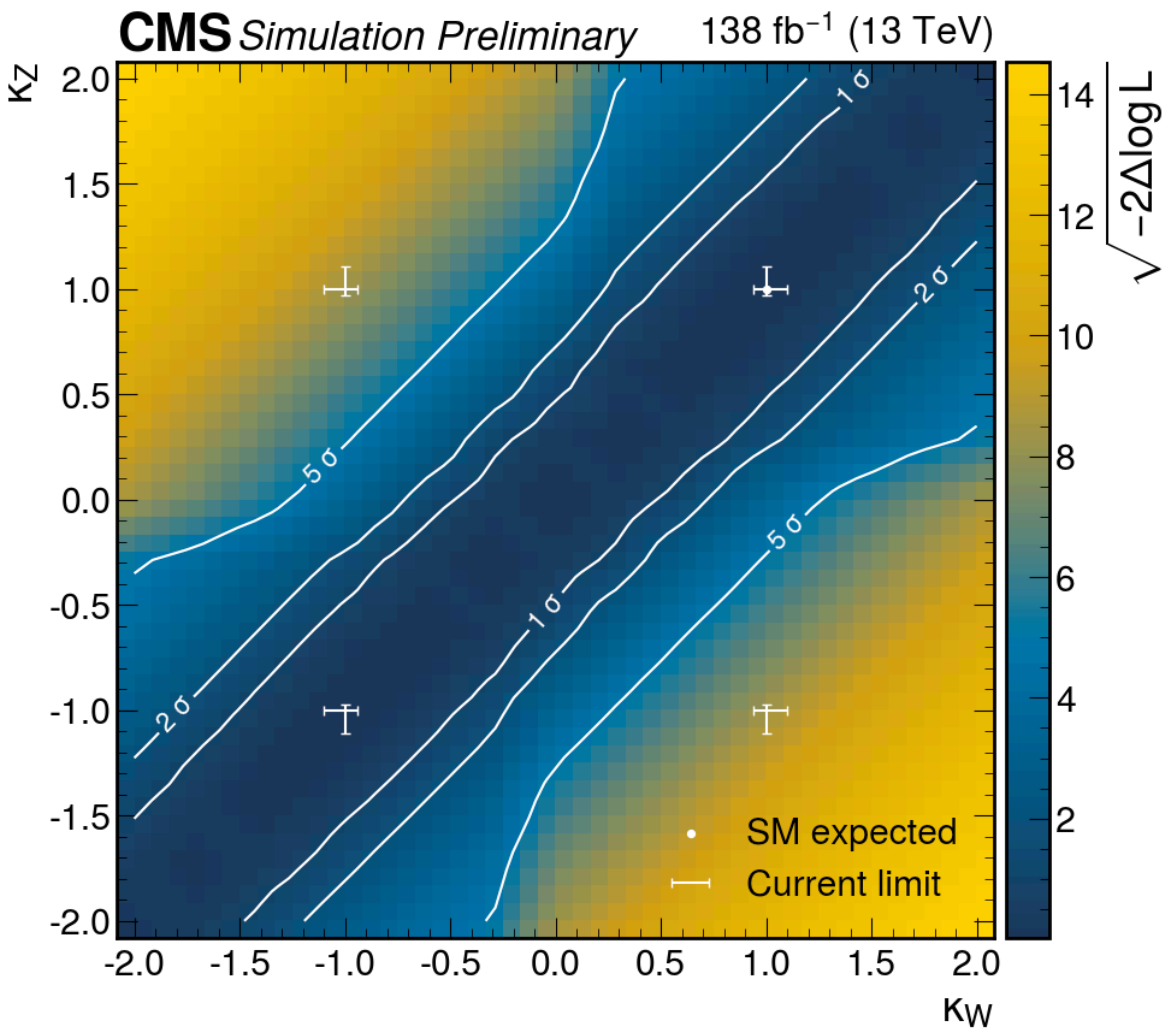
Expected signal (λ_{WZ} = -1): 366 ± 2.9
Predicted background: 120 ± 16.1 ± 15.3
stat. *syst.*

*https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/

UC San Diego

Detailed Results (cont.)

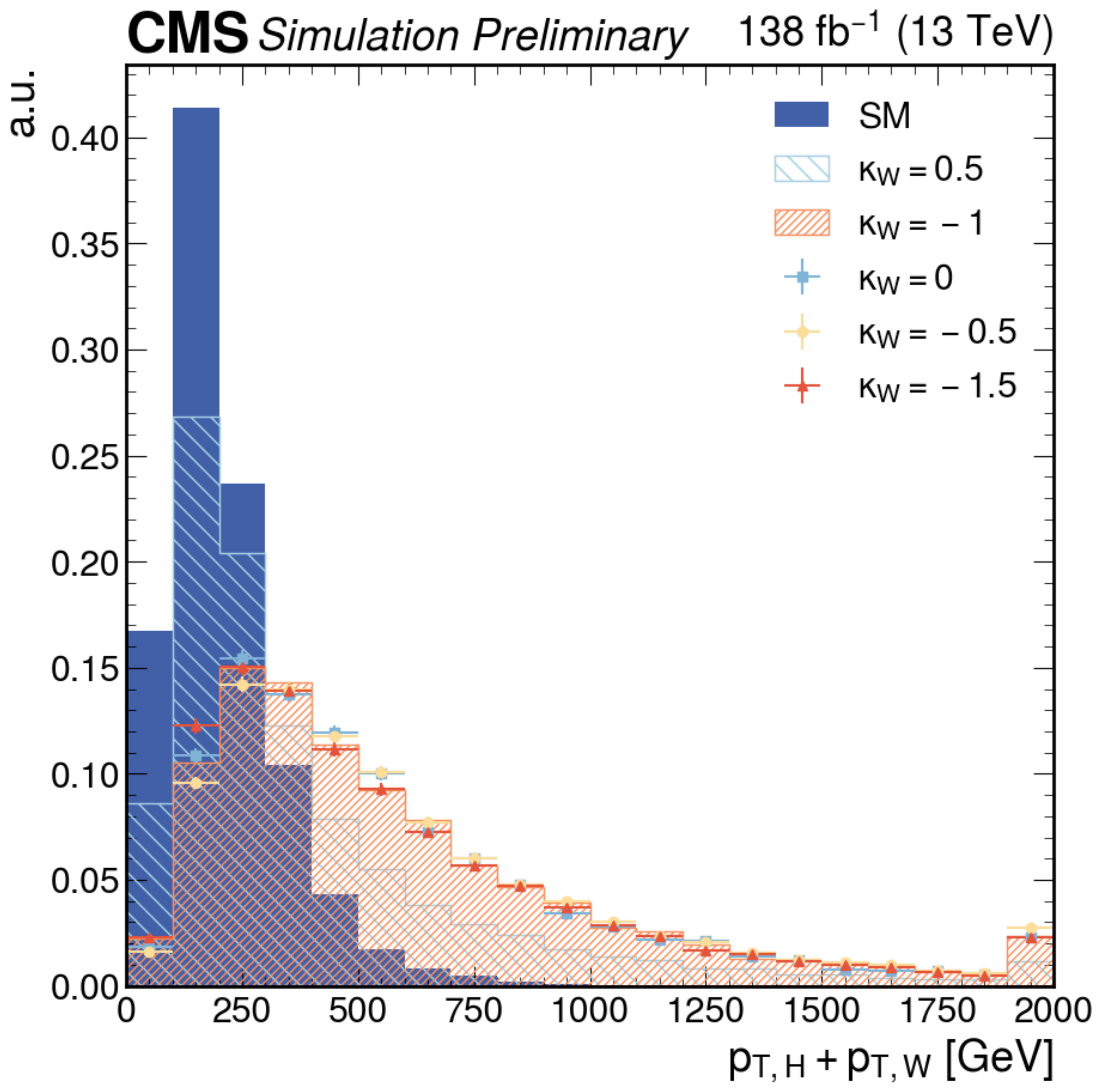
- Used MadGraph reweighing to scan many κ_W, κ_Z values
- Interpolated exclusion limits plotted on z-axis
- Current best limits on $|\kappa_V|$ are plotted as capped "error bars" (represent 1D limits, not 2D errors)
  - $|\kappa_W| = 1.02 \pm 0.08$, $|\kappa_Z| = 1.04 \pm 0.07$
- Contours show $\sigma = 1, 2, 5$ exclusion boundaries
  - This shows we **can exclude $\lambda_{WZ} < 0$** when considered alongside current best limits



UC San Diego

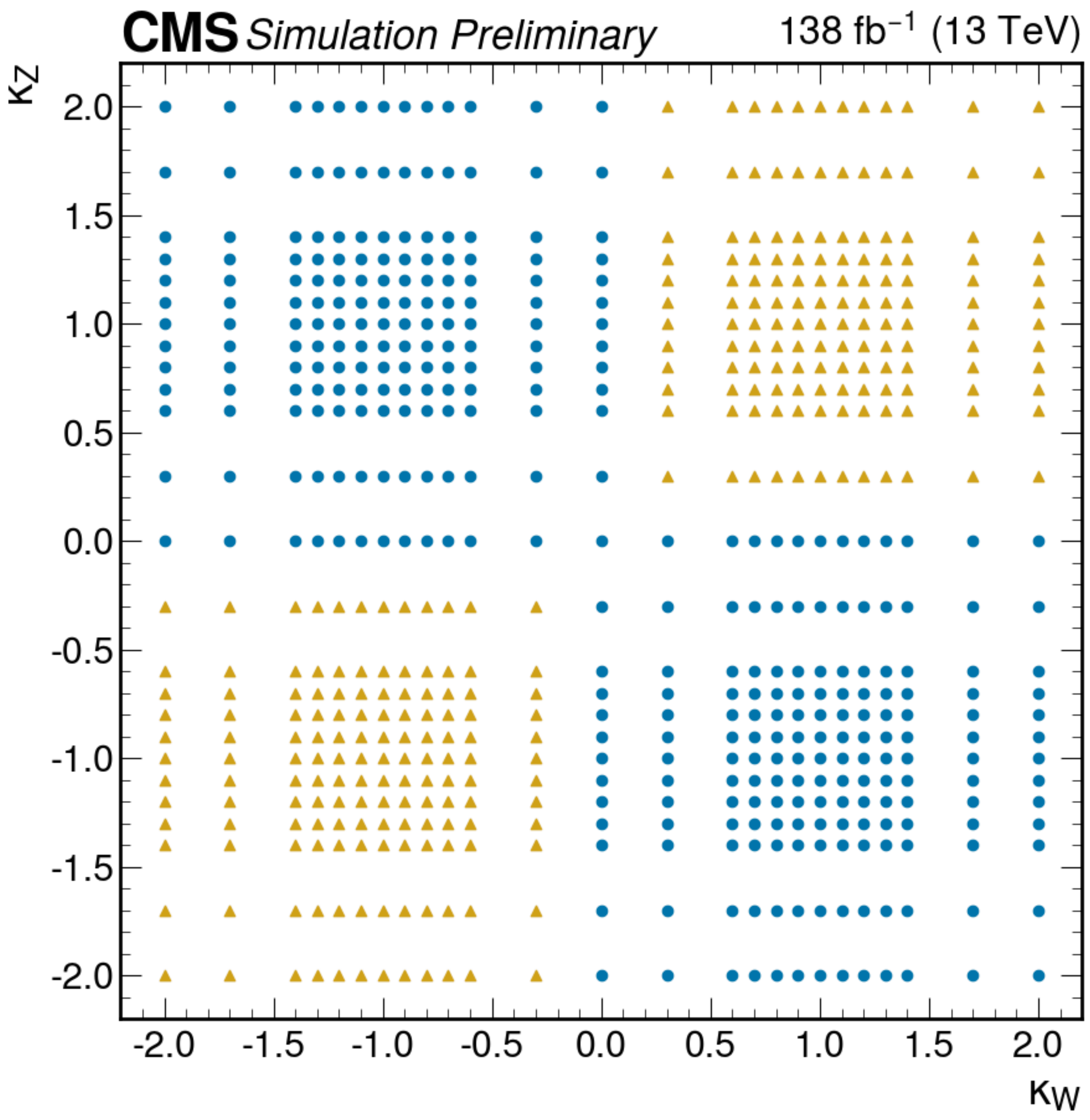
Other λ_{WZ} values

- Generated 10k events for various κ_W values
- Comparing kinematics at LHE level
  - Not much difference across a fairly large range of κ_W values
  - ⇒ Acceptance ~consistent for $\kappa_W = -1 \pm \epsilon$

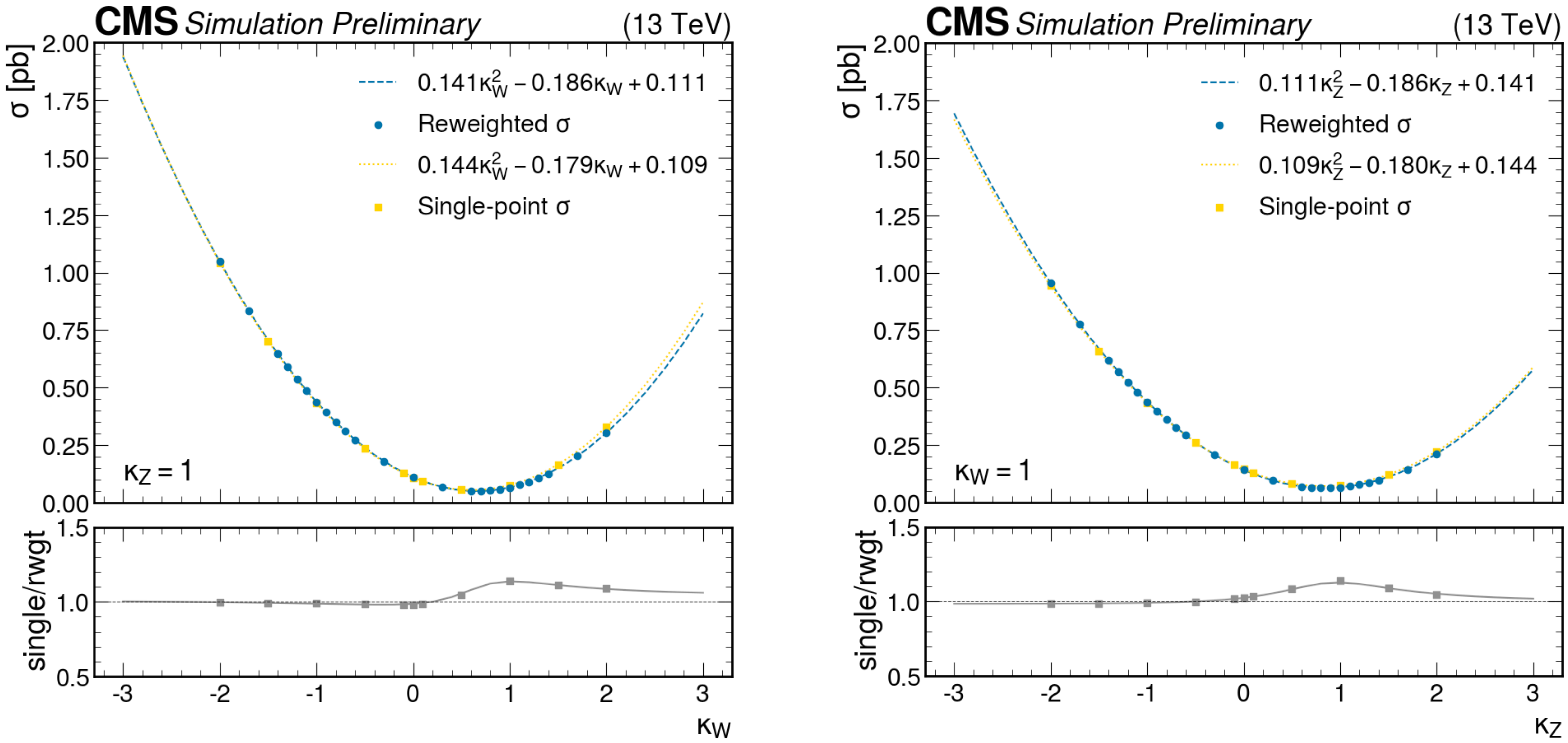


λ_{WZ} Scan

- Generated two signal samples:
  - $\lambda_{WZ} \leq 0$ sample
    - Reweighted around (κ_W = -1, κ_Z = +1)
  - $\lambda_{WZ} > 0$ sample
    - Reweighted around (κ_W = +1, κ_Z = +1)
  - Used PKU reweighting model
  - Full Run 2 samples
  - 100k events per NanoAODv9 “year”



New Signal Samples: Validation

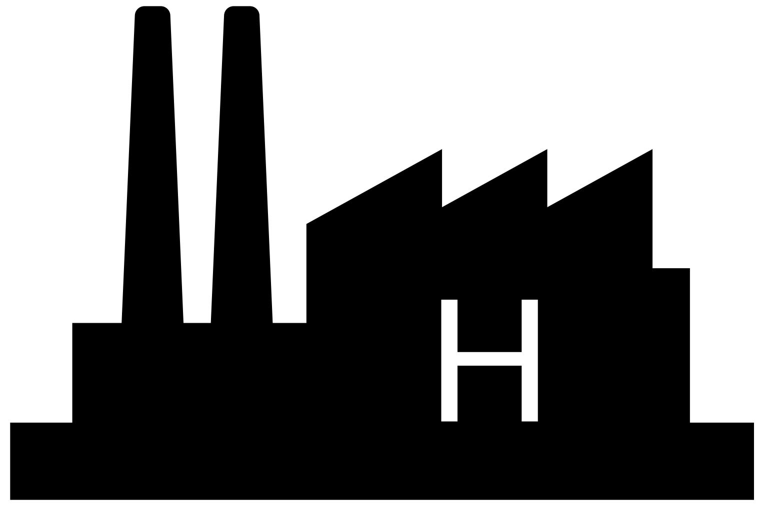


UC San Diego

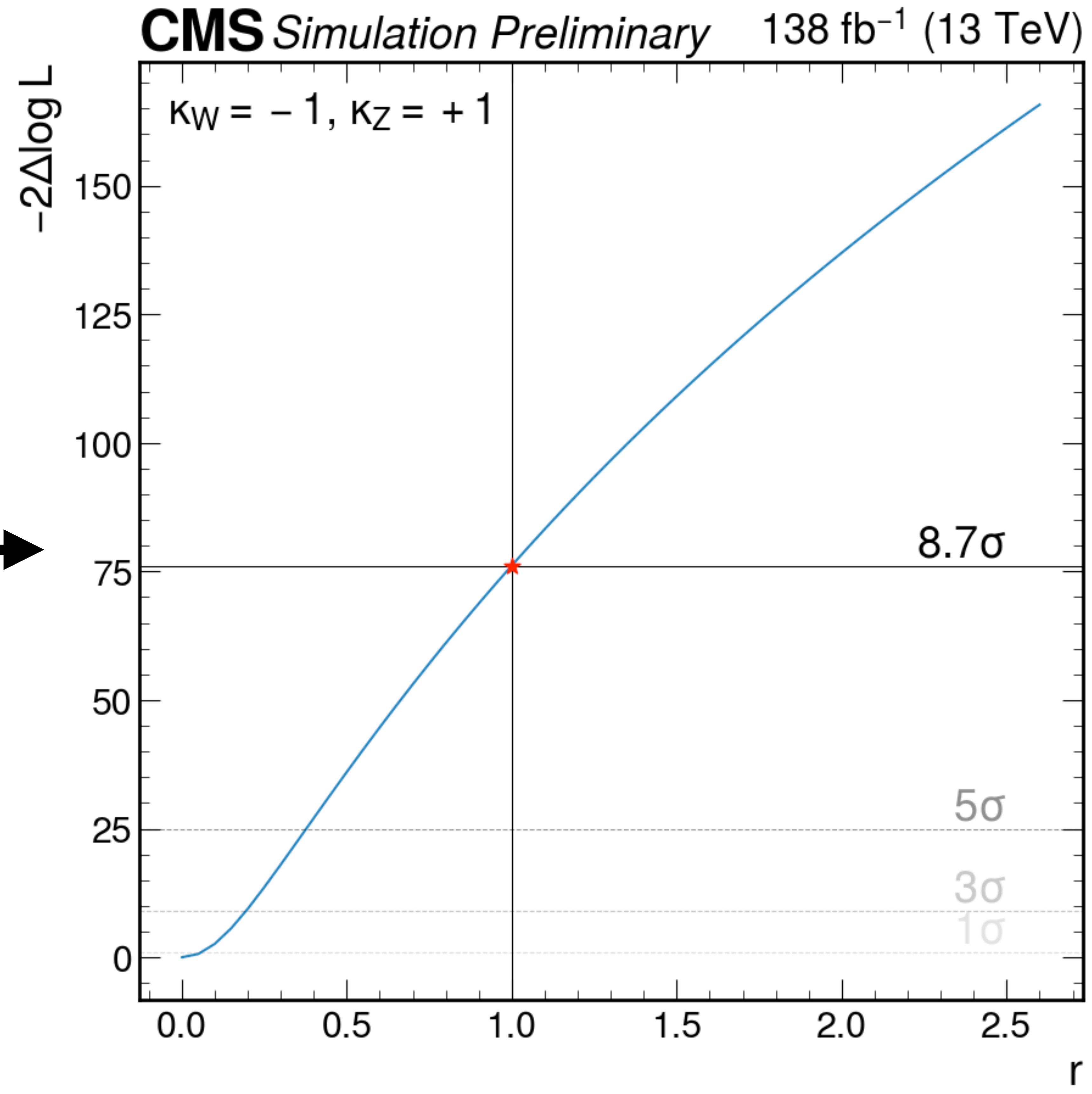
HiggsCombine Settings

```
imax 1 number of channels
jmax 1 number of backgrounds
kmax 19 number of nuisance parameters
------------------------------------------------
bin                              bin1
observation                       128
------------------------------------------------
bin                              bin1        bin1
process                     VBSWH_mkW    TotalBkg
process                             0           1
rate                           413.34      127.92
------------------------------------------------
abcd_syst     lnN                   -       1.084
abcd_stat     lnN                   -       1.133
pdf_vars      lnN               1.022           -
muF_scale     lnN               1.178           -
isr_weights   lnN               1.001           -
fsr_weights   lnN               1.015           -
pu_rwgt       lnN               1.002           -
L1_prefire    lnN               1.010           -
hlt_sfs       lnN               1.008           -
mc_stat       lnN               1.022           -
lep_id        lnN               1.015           -
elec_reco     lnN               1.003           -
muon_iso      lnN               1.000           -
xbb_sfs       lnN               1.057           -
btag_sfs      lnN               1.003           -
met_unc       lnN               1.003           -
jes           lnN               1.066           -
jer           lnN               1.008           -
lumi          lnN               1.025           -
```

scan_kW_X_kZ_Y.dat

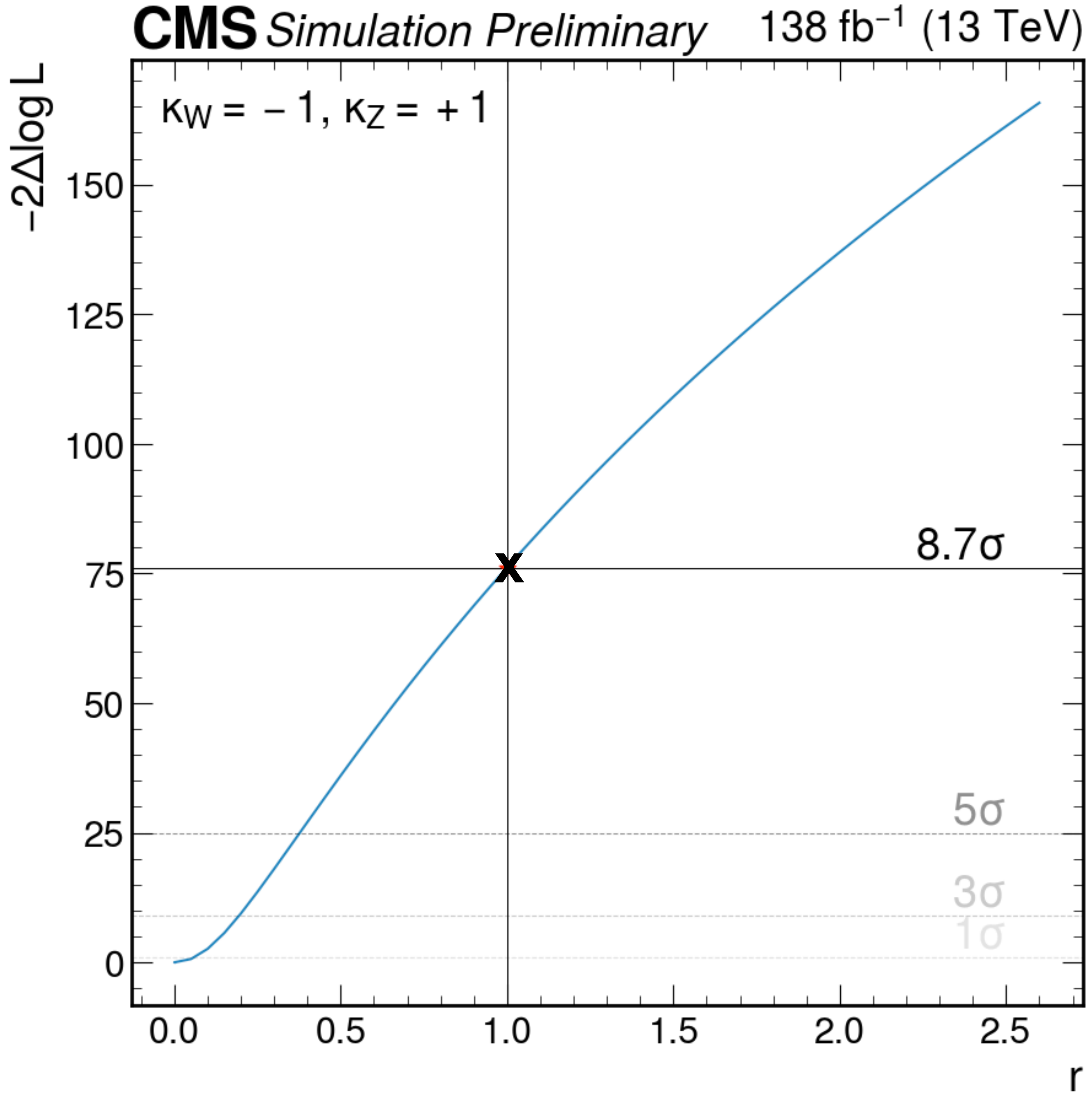



```
combine -M MultiDimFit -d scan_kW_X_kZ_Y.root
    -m 125 -t -1
    --expectSignal=0
    --setParameters r_VBSWH_mKW=0
    --setParameterRanges r_VBSWH_mKW=0.0,2.0
    --saveNLL
    --algo grid
    --points 101
    --rMin 0 --rMax 5
    --alignEdges 1
```

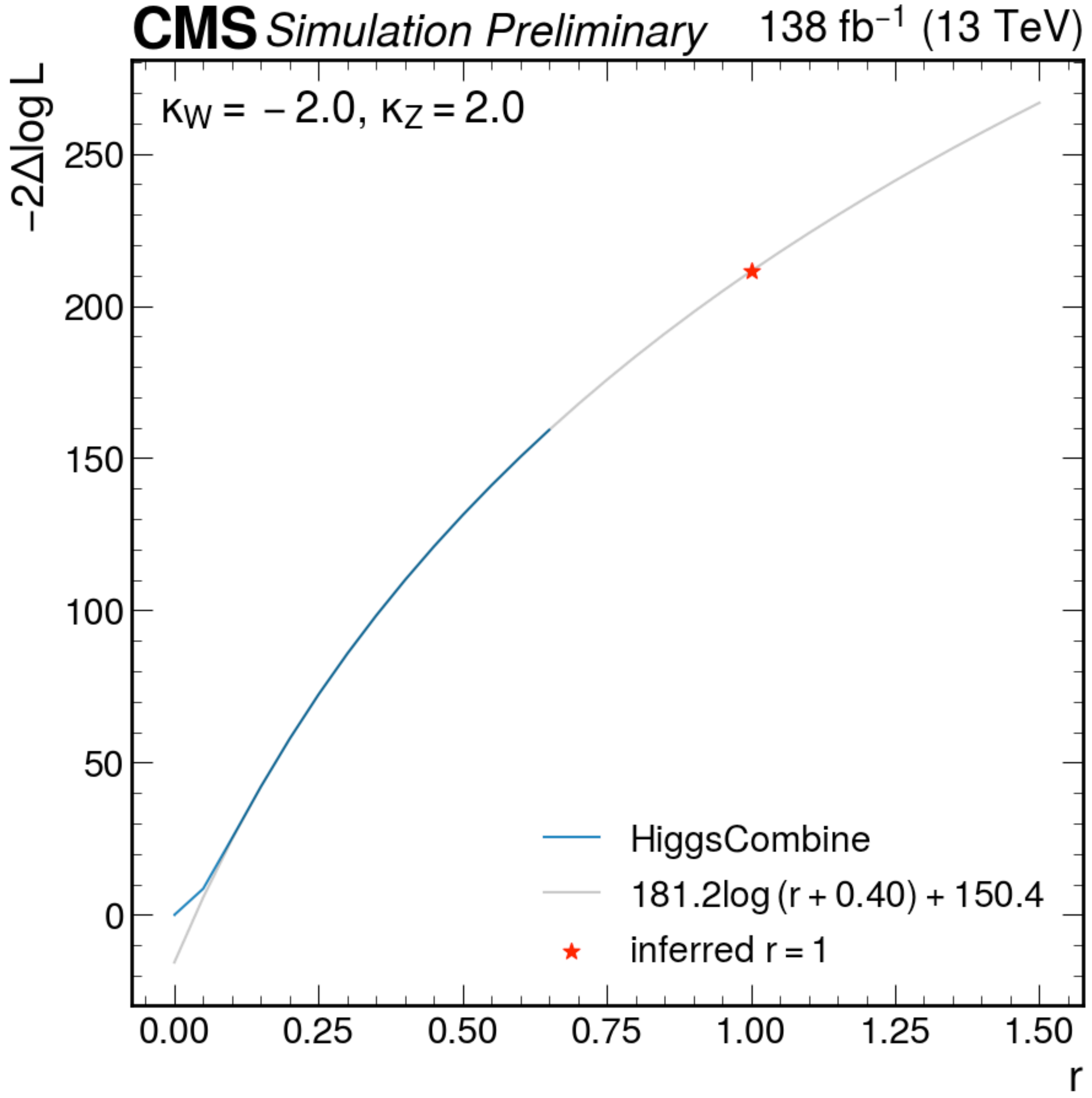


Repeat for each point $\kappa_W = X$, $\kappa_Z = Y$

Extrapolated Points



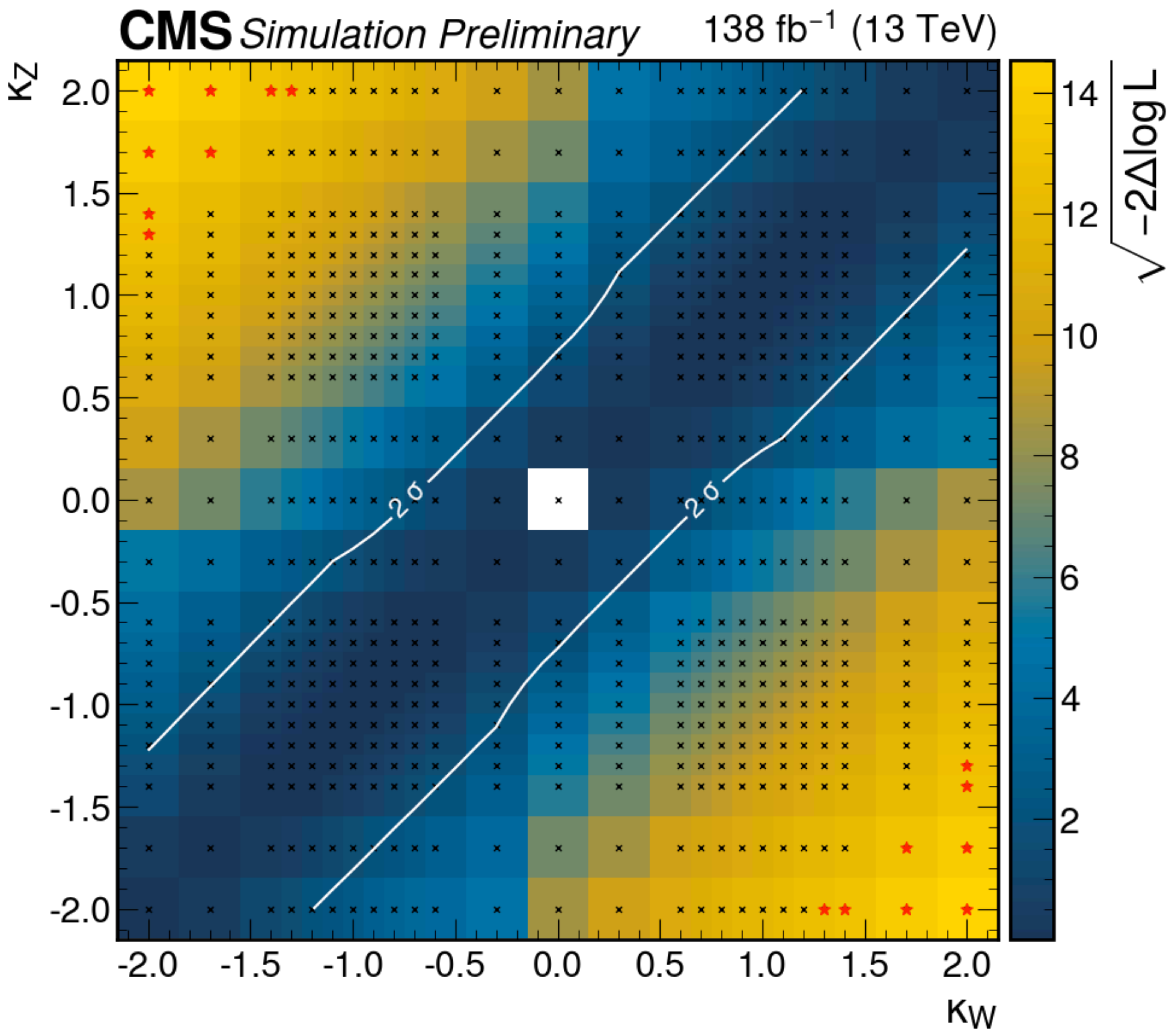
Take σ exclusion of r = 1



Infer σ exclusion of r = 1

Collected Results

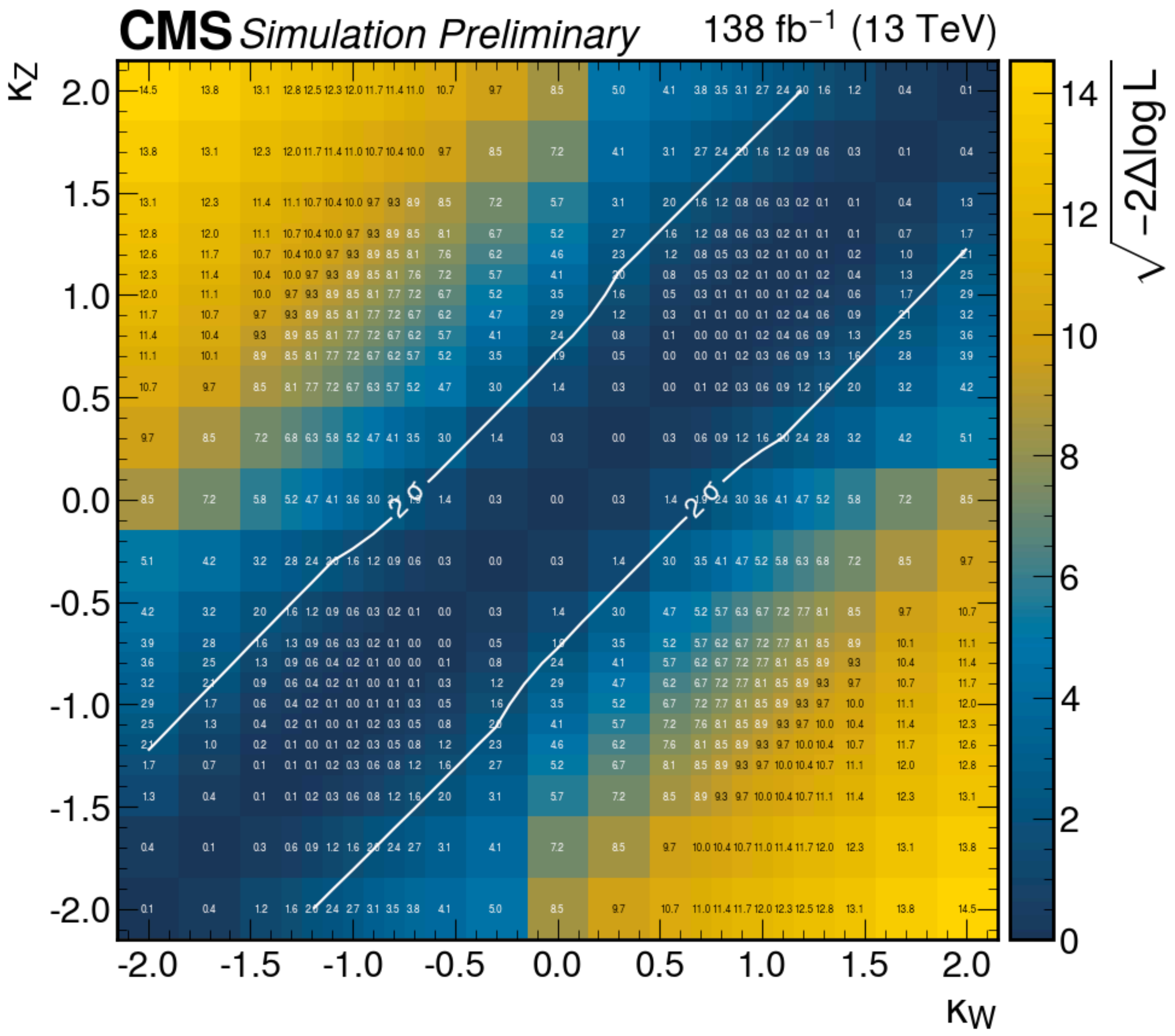
- Bins centered on scanned κ_W, κ_Z points
- Exclusion limit plotted on z-axis
- Contour roughly shows $\sigma = 2$ boundary
  - Simplistically derived by Matplotlib
- Black **x**'s taken directly from HiggsCombine plot
- Red ★'s inferred from plot



UC San Diego

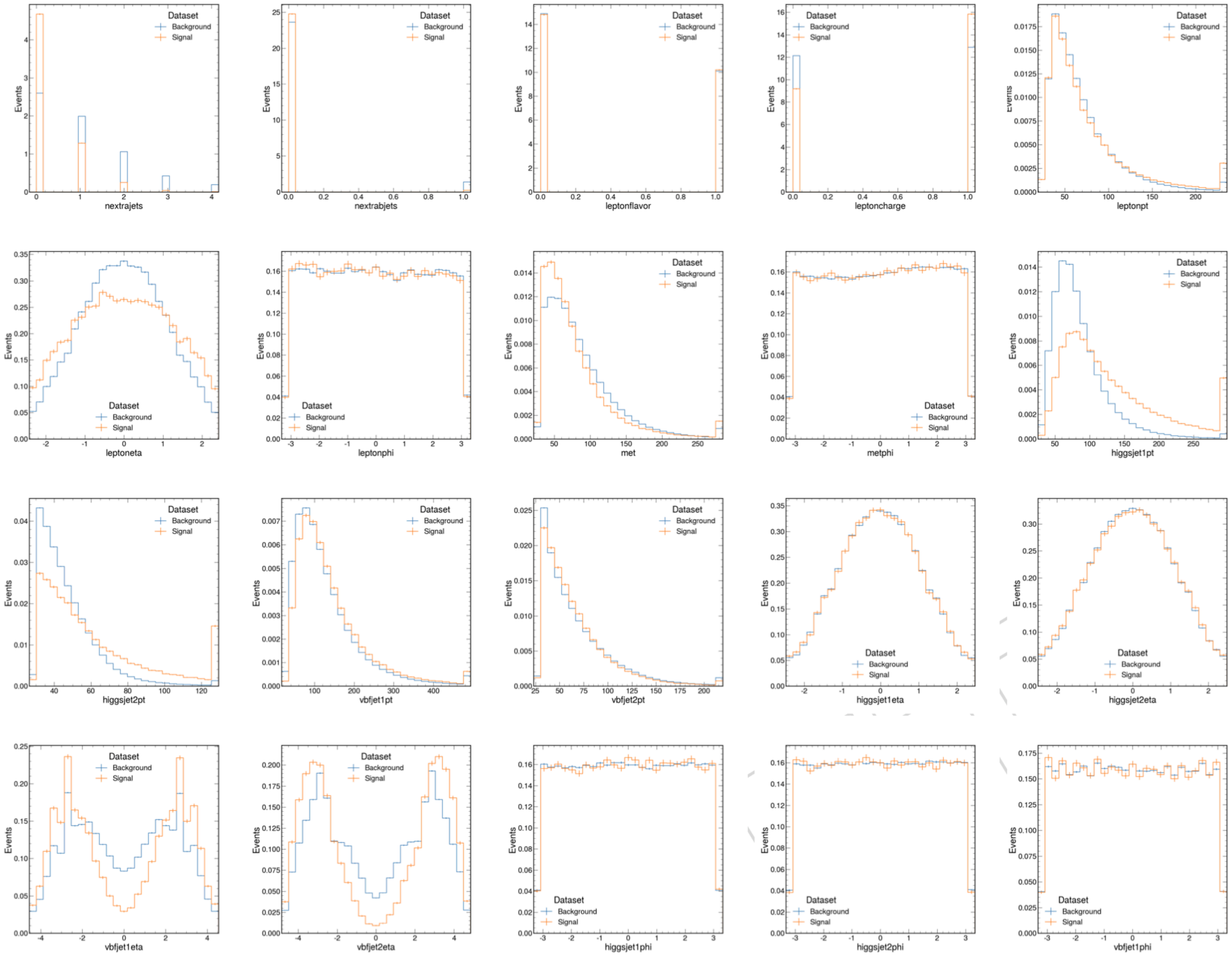
Collected Results

- Bins centered on scanned κ_W, κ_Z points
- Exclusion limit plotted on z-axis
- Contour roughly shows σ = 2 boundary
    - Simplistically derived by Matplotlib
- Discontinuities do not affect contours
    - Caused by some failure in the reweighting
    - Cross sections are reweighted properly, acceptance is not
    - Smoothed out via interpolation

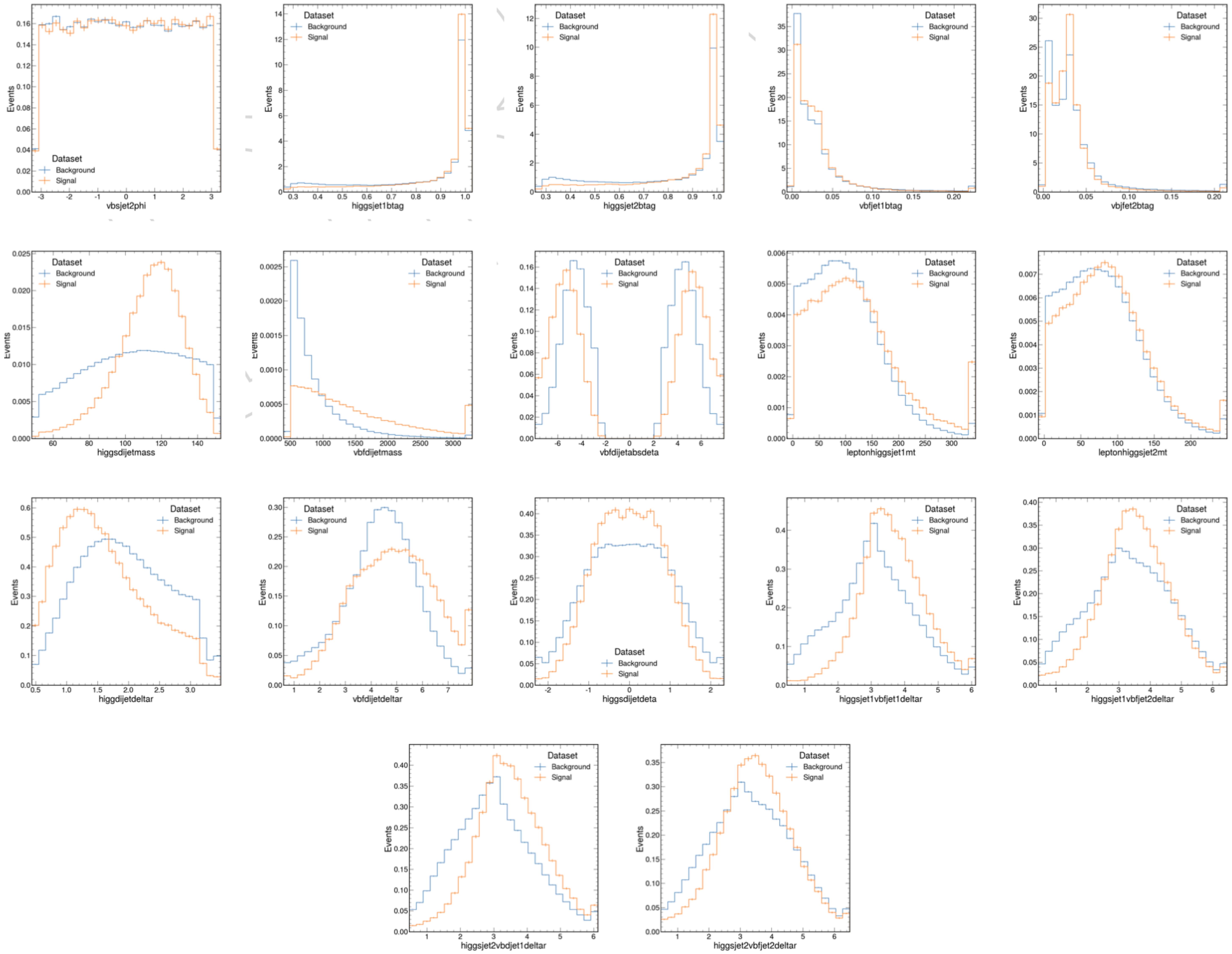


Backup (resolved)

BDT Input Variables



BDT Input Variables (cont.)



BDT Training

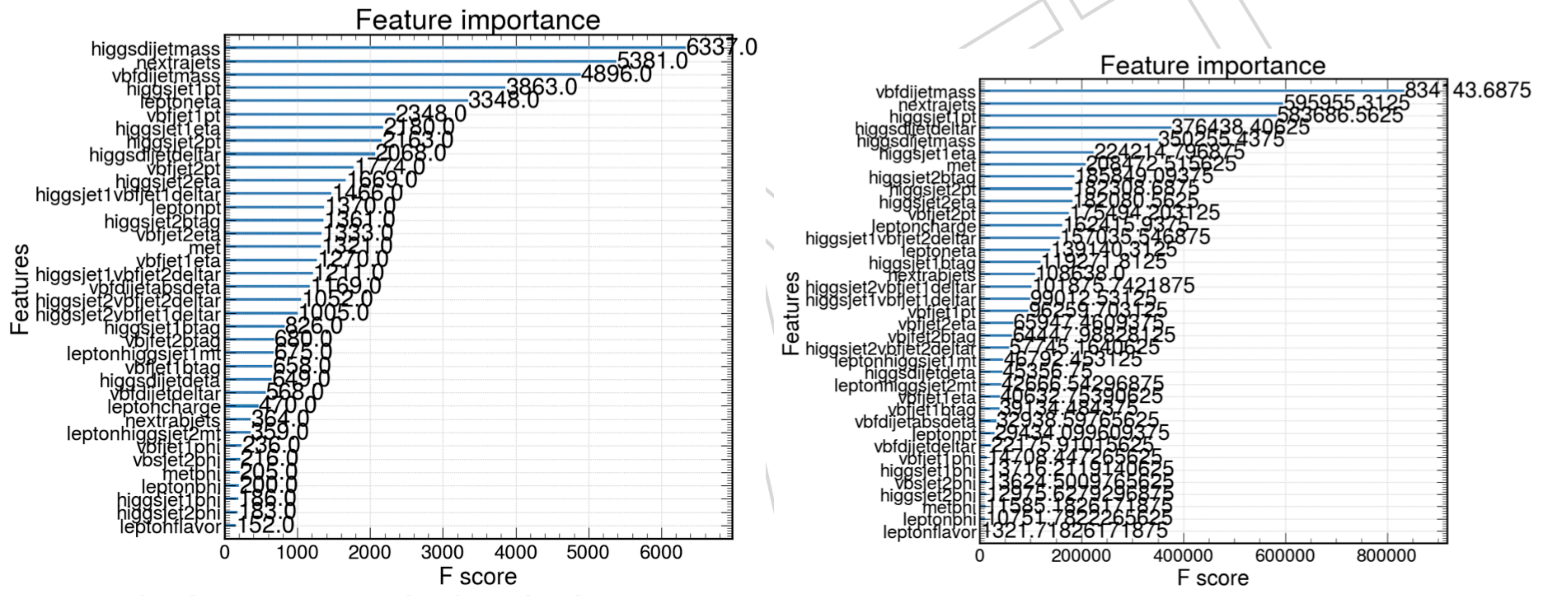


Figure 4: Feature importance as measured with the weight metric (left) and gain metric (right).

Weight metric: the number of times that feature is used to split the data
Gain metric: the average increase in the objective function for all splits based on that feature

BDT settings

Learning Rate	0.01
Max Depth	6
Early stopping rounds	1

- Used 90/10 train/test split
- Set weights = abs(weights), but affects < 1% of events