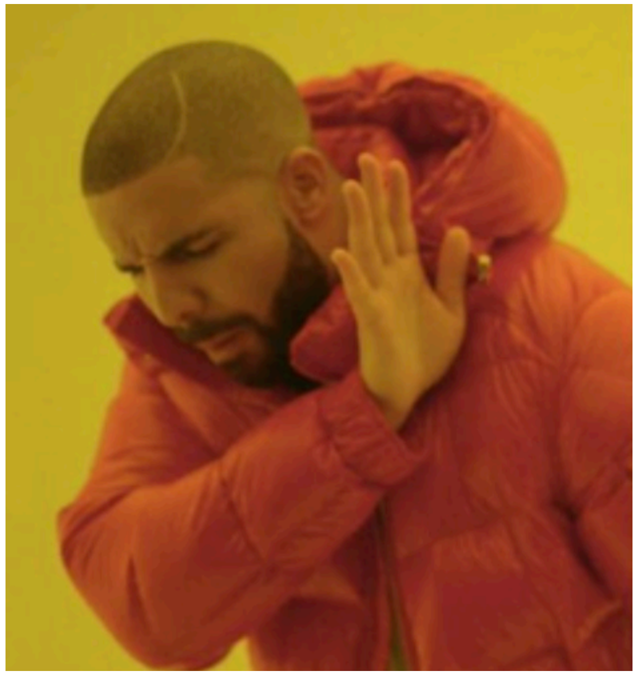
Accessing entanglement and suppressing background in semileptonic $H \to WW^*$

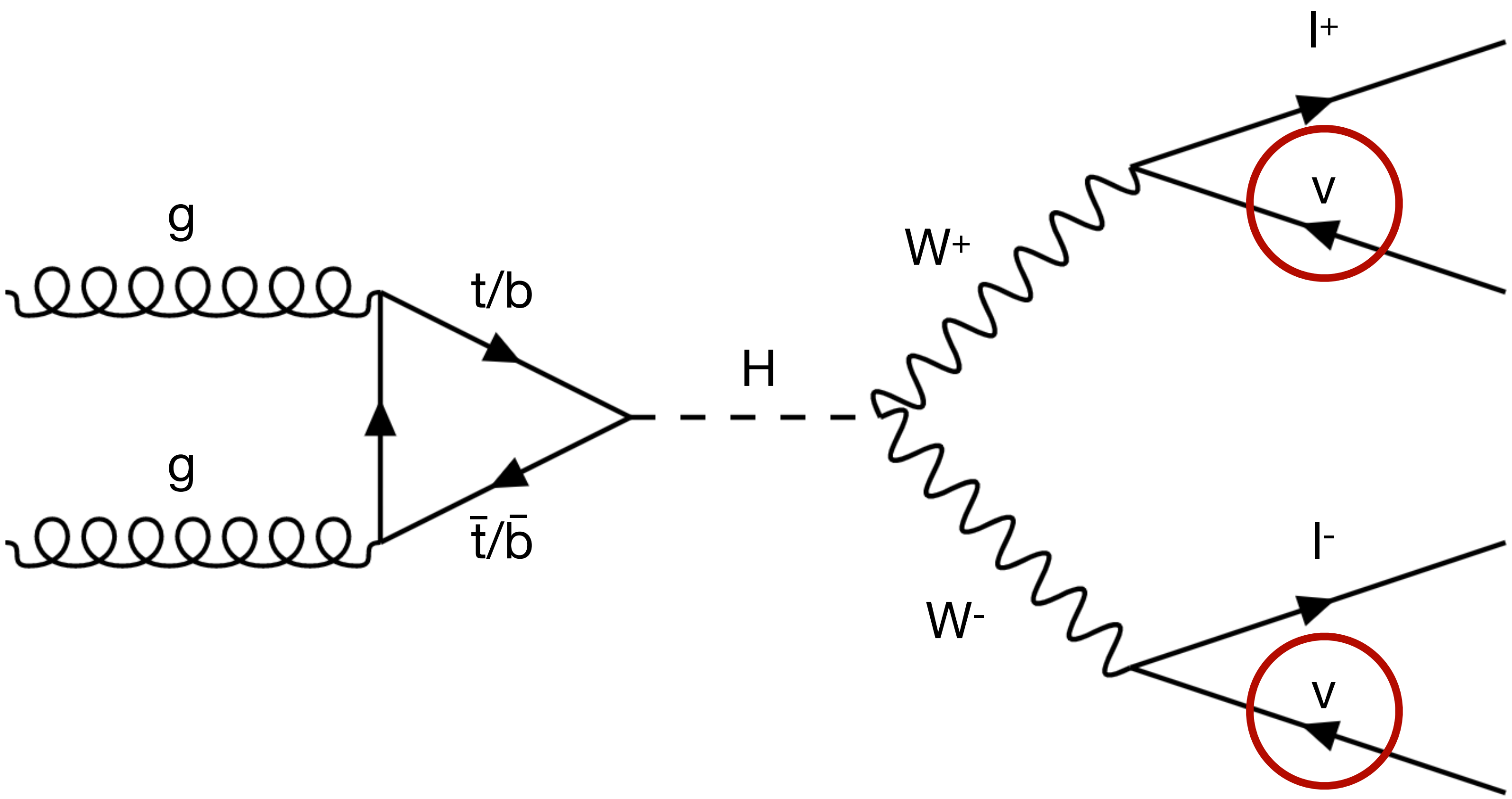
Federica Fabbri, Jay Howarth, **Théo Maurin**

University of Glasgow
UKRI Science and Technology Facilities Council
The Royal Society

Isolating Signal

- **So, why not use $H \rightarrow WW^* \rightarrow l\nu l\nu$ for Quantum information measurements?**

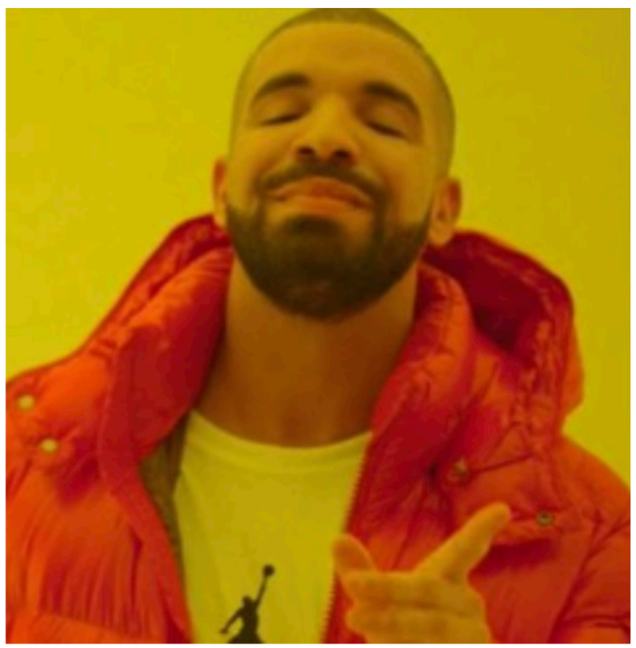


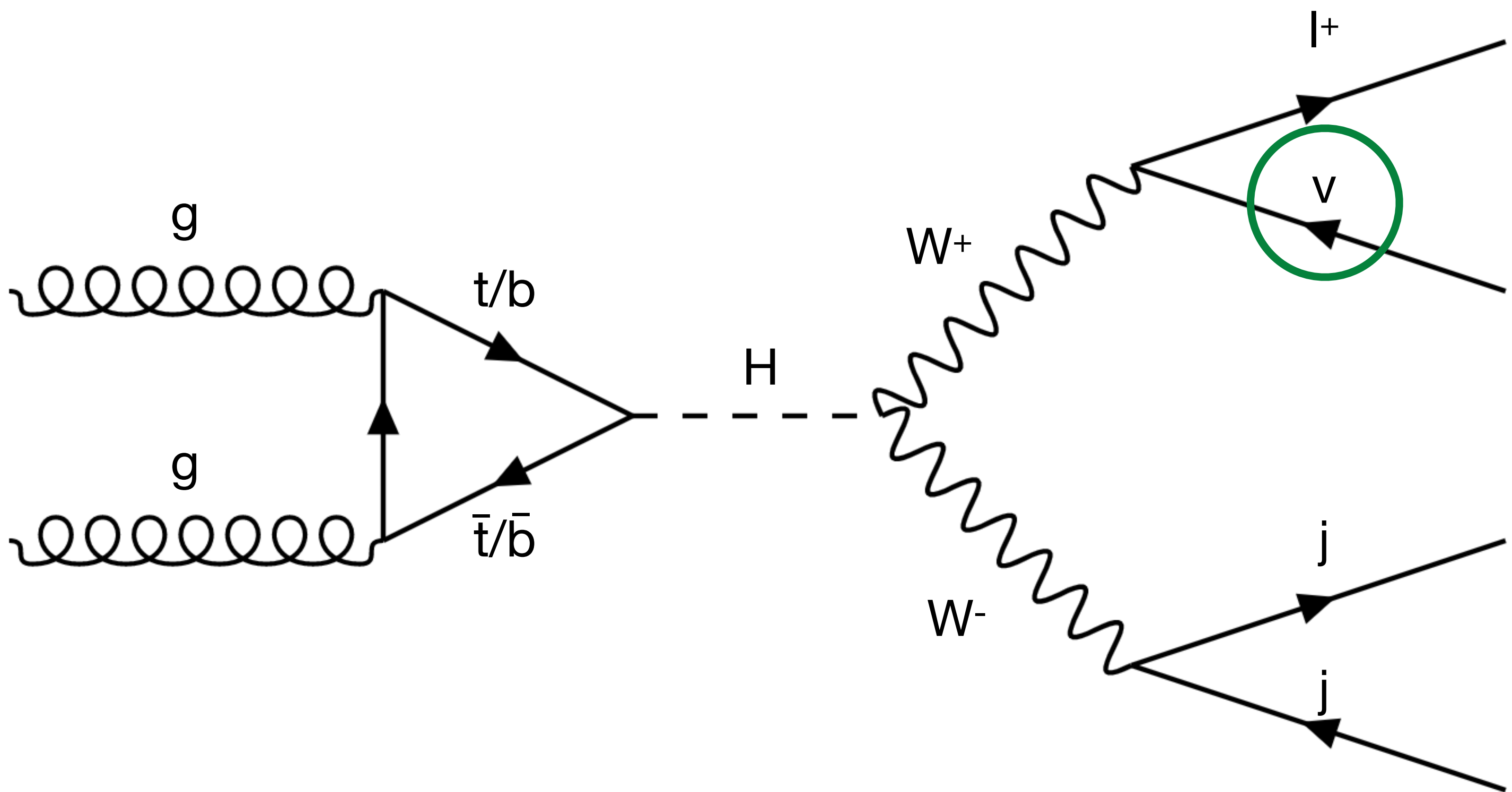


- **Presence of two neutrinos** (and lack of additional mass constraints due to off-shell W) **makes it extremely hard to reconstruct the Higgs.**

Isolating Signal

- **Using the $H\rightarrow WW^{*}\rightarrow jjl\nu$ channel we can reconstruct the Higgs:**



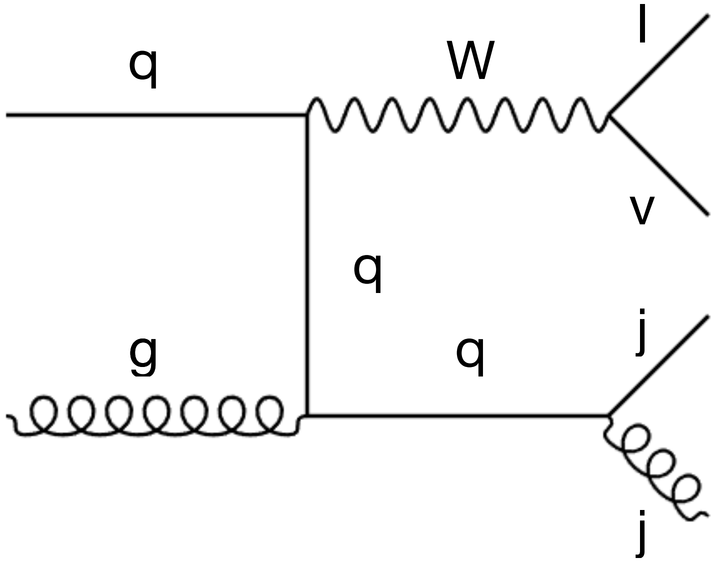


- **But there are two challenges:**
  1. Dealing with overwhelming W+jets background.
  2. Accessing the spin information of the hadronic W.

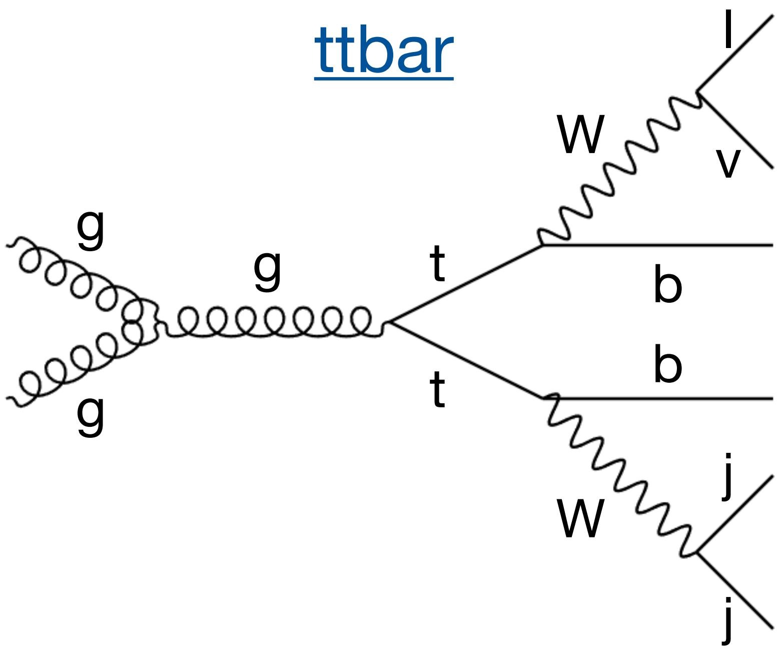
Isolating Signal

- **The backgrounds**

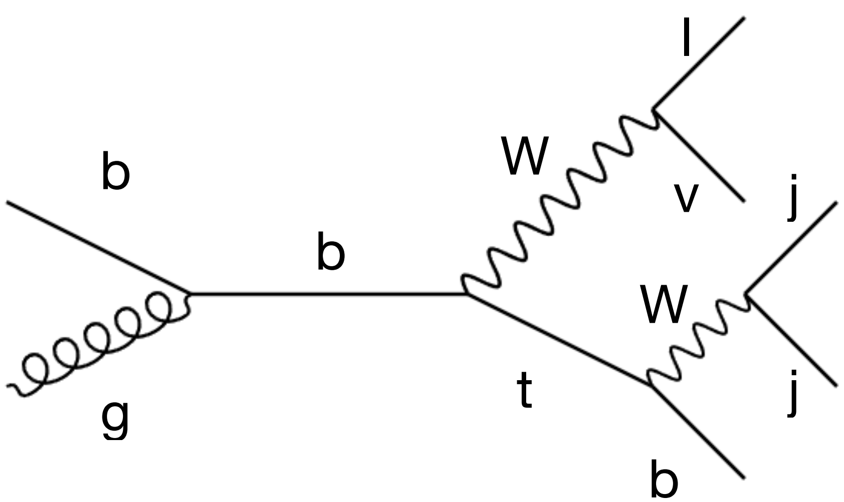
W + jets



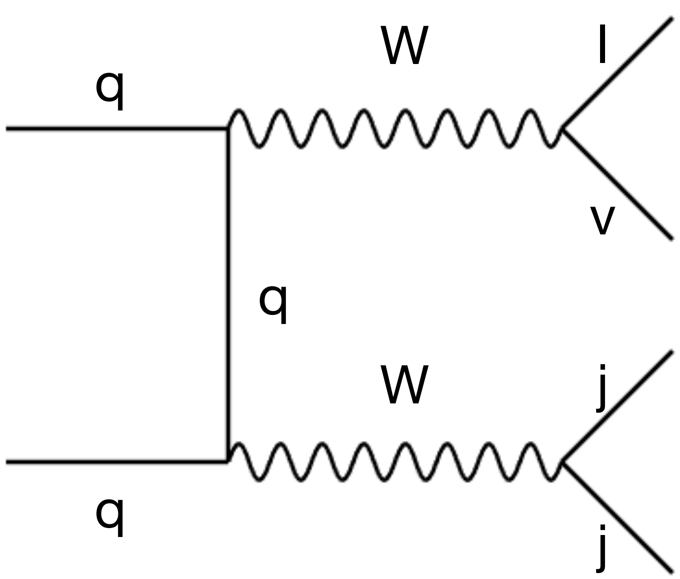
ttbar



t+W



WW



Accessing Spin Info

- **W spin information can be accessed either using charged leptons or down-type quarks:**

"Spin Analysing Power"
degree to which a particle caries parent spin info, 1 = fully, 0 = none, -1 = 'anti' fully

	ℓ	u, c	d, s
LO	1.000	-0.310	1.000
NLO	0.998	-0.310	0.930

- **If we can identify the down-type jet in a hadronic W decay, we can access the spin information.**

- **W→cs allows us to do this because we can tag the c-flavoured jet and then take the other jet that pairs to make the W mass as our s-jet** (spin analyser)**.**

Monte Carlo Simulations

- **We used the following MC generators to simulate signal and background processes:**

  ➡ **Higgs** (gg fusion)**: Powheg + Pythia8**
  ➡ **W+jets: Powheg + Pythia8**
  ➡ **ttbar: Powheg + Pythia8**
  ➡ **Diboson: Sherpa** (NLO)

- **We generate about 900k signal events and ~1 mil background events for each process.**

Simulating Collider Conditions

- Cuts and reconstruction/trigger efficiencies are applied to jets and leptons to simulate detector effects:

$$p_t > 25 \text{ GeV} \qquad |\eta| < 2.5$$

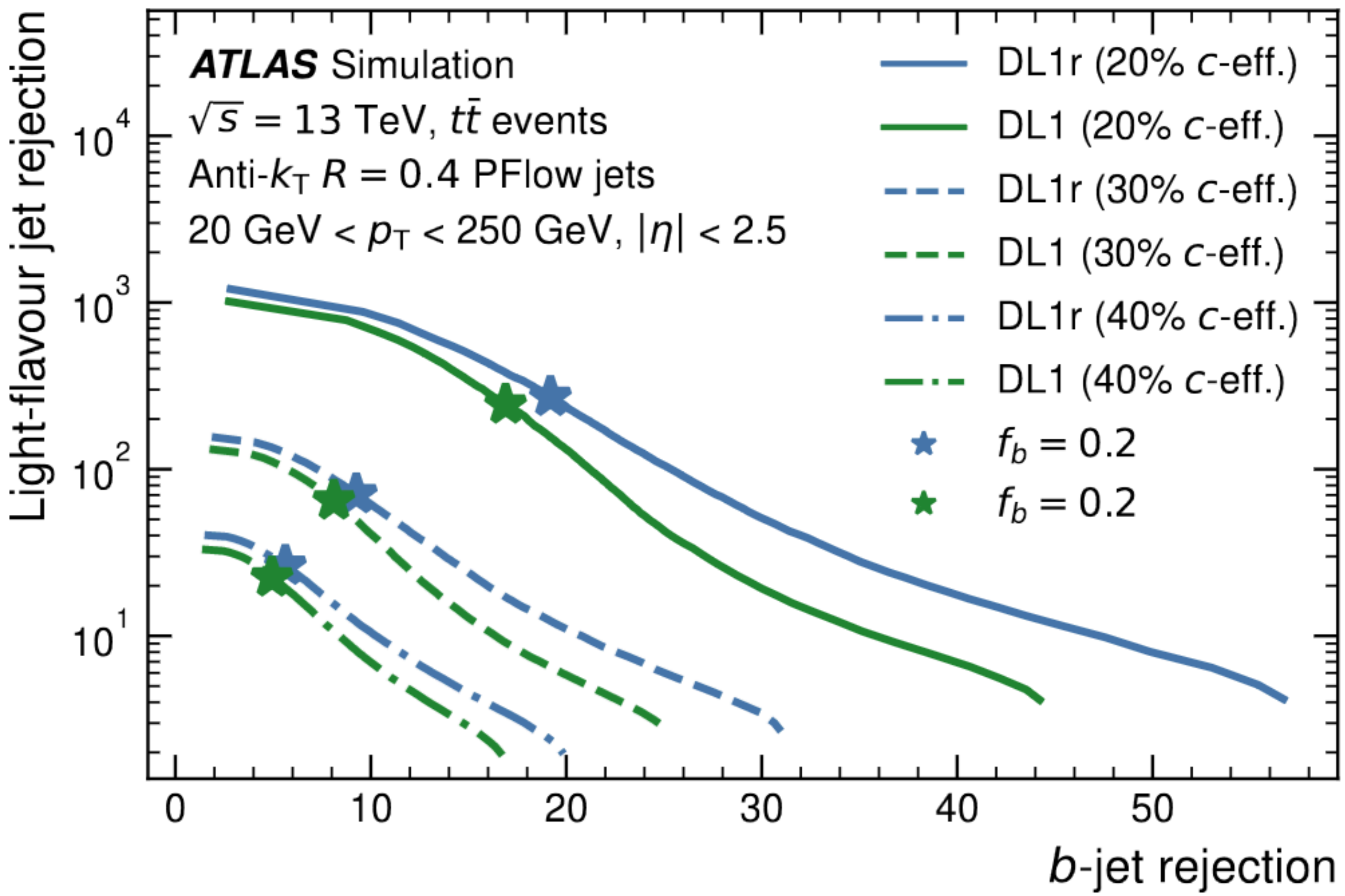
- Jets are 'tagged' to approximate realistic c- and b-tagging efficiencies.

Efficiency type	Efficiencies (%)	
	c-tagger	b-tagger
ϵ_b	14	77
ϵ_c	40	20
ϵ_l	3.3	0.8

- All of these together approximate very roughly the reco level.

Charm Tagging

- **Key question is how well can we tag charm jets experimentally?**



Taken from https://arxiv.org/pdf/2211.16345.pdf

Reconstructing the Higgs

- **Using $W \rightarrow cs$ decays requires the hadronic W to be on-shell and therefore the leptonic to be off-shell.**

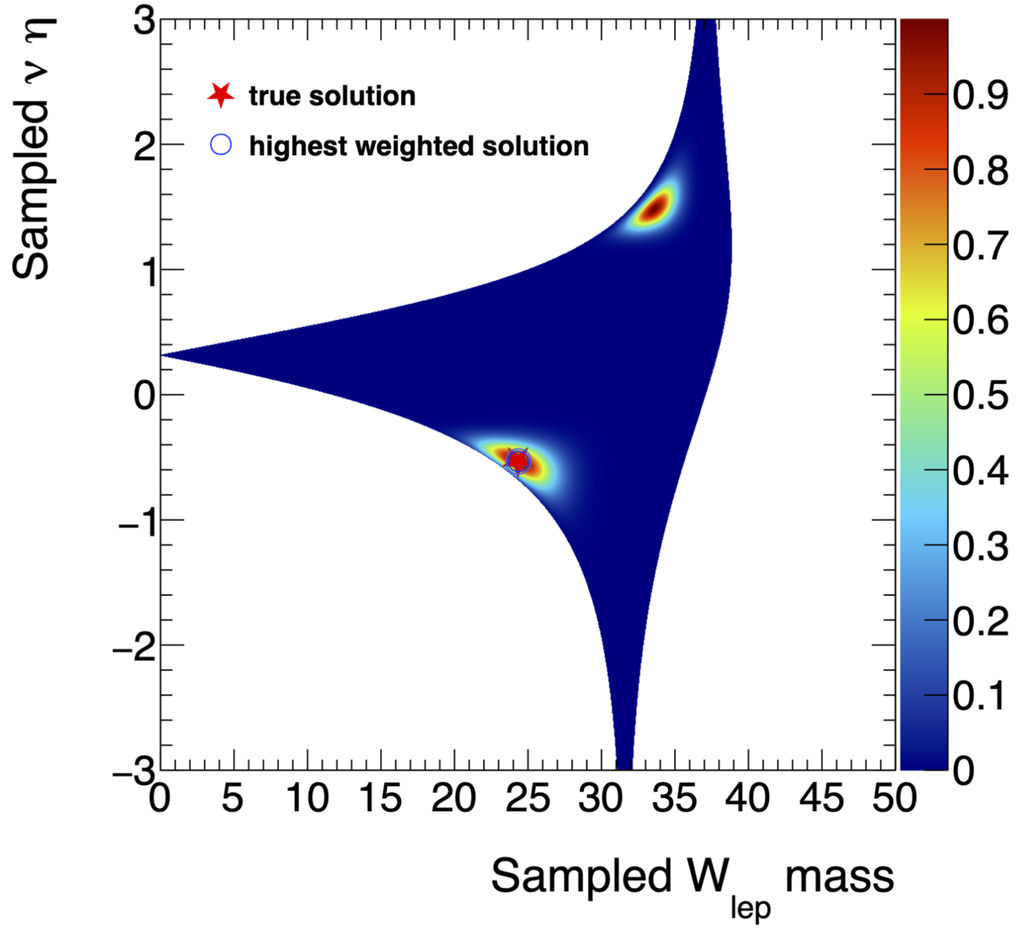
- **We have two unknowns remaining:**
  1. The long. component of the neutrino momentum
  2. The inv. mass of the off-shell W boson (m_W < 80 GeV)

- **This is similar to a problem in $t\bar{t}$ final states, where we use a tool called 'Neutrino Weighting' to reconstruct events by integrating over missing kinematics.**

- **In this case, we integrate over 1. and 2.**

Reconstructing the Higgs

- **We reconstruct many Higgs each event under different assumptions of m_{W^*} and η_ν.**



- **Each solution yields a weight based on how well the reconstructed ν agrees with the observed MET** (missing transverse momentum in the event)**.**

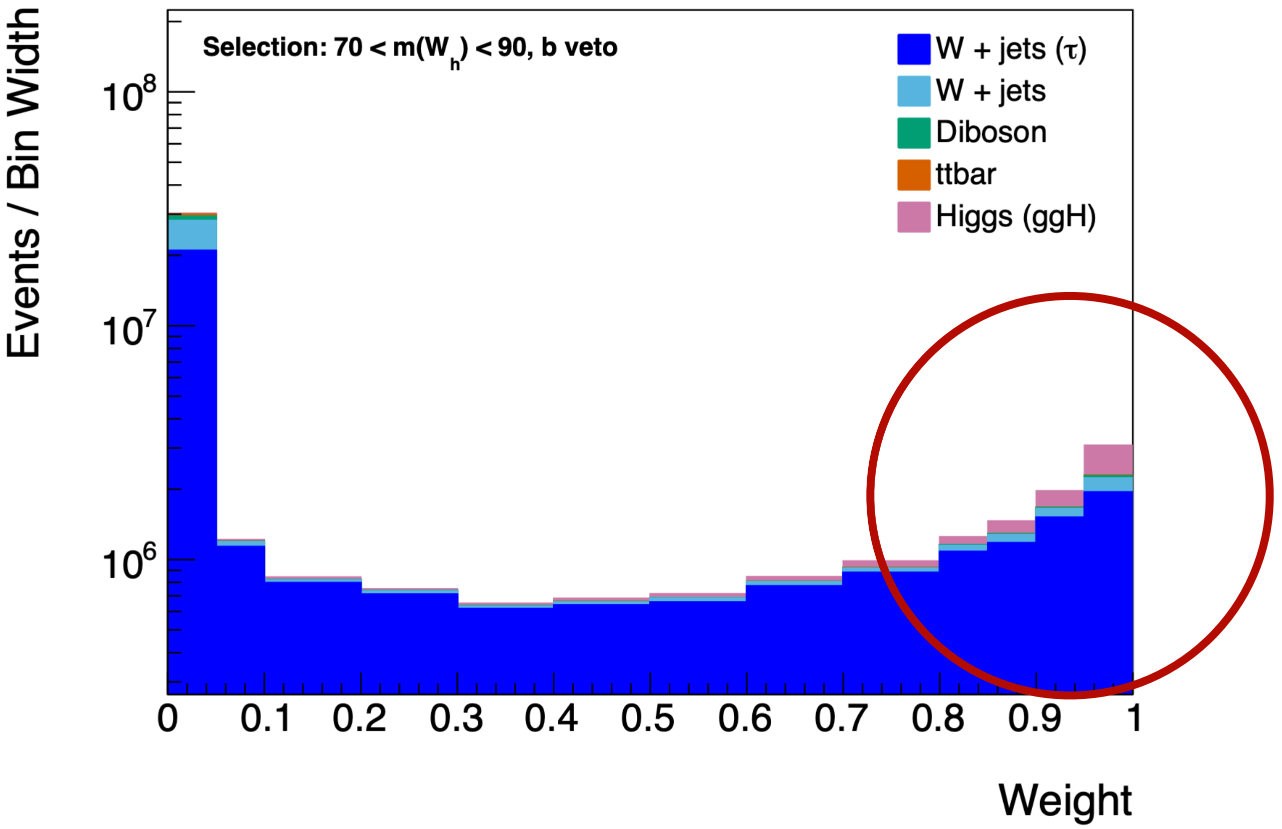
- **The solution with the highest weight is taken as the correct one.**

The W+jets background.

- **Using W→cs decays requires the hadronic W to be on-shell and therefore the leptonic to be off-shell.**

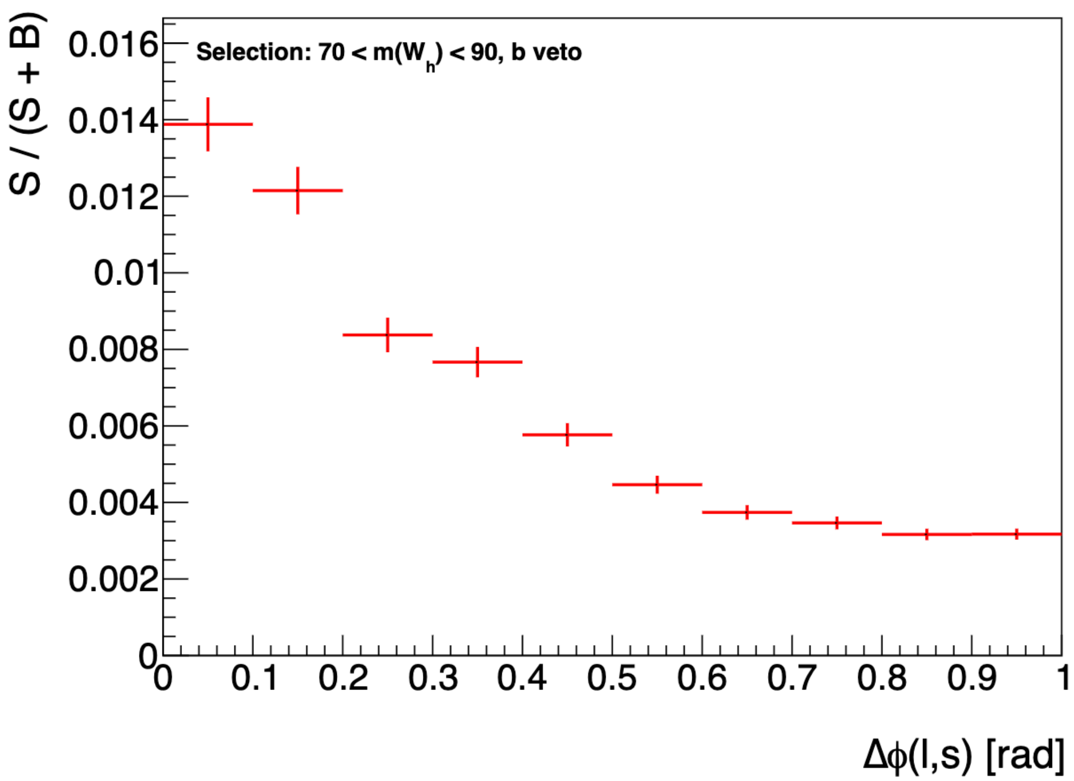
- **This is a bonus because the primary background** (W→lνjj) **has an on-shell leptonic W.**

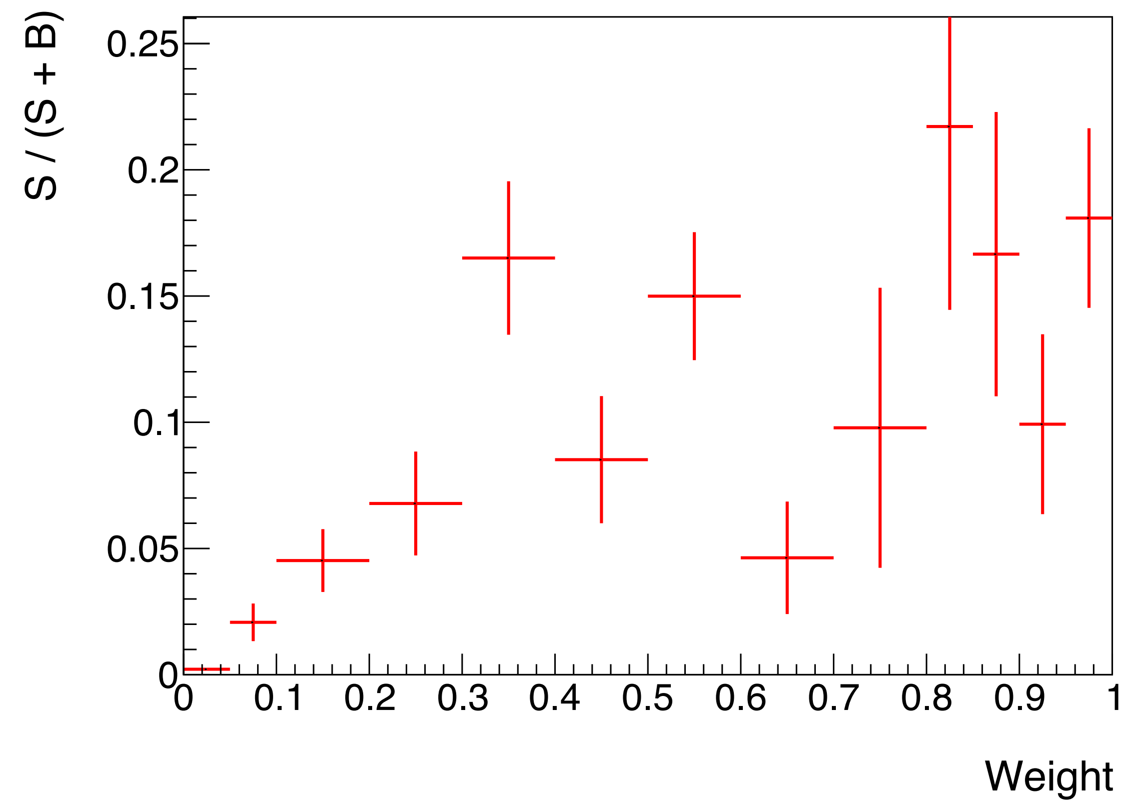
- **The weight from NW also acts similarly to an MVA classifier** (signal events are more likely to result in higher weights than background events).



Cuts to remove background

- **We identified two parameters to cut on to enhance signal: $\Delta\Phi_{ls}$** (difference in Φ of the spin analysing pair (l,s)) **and the weight from the NW.**

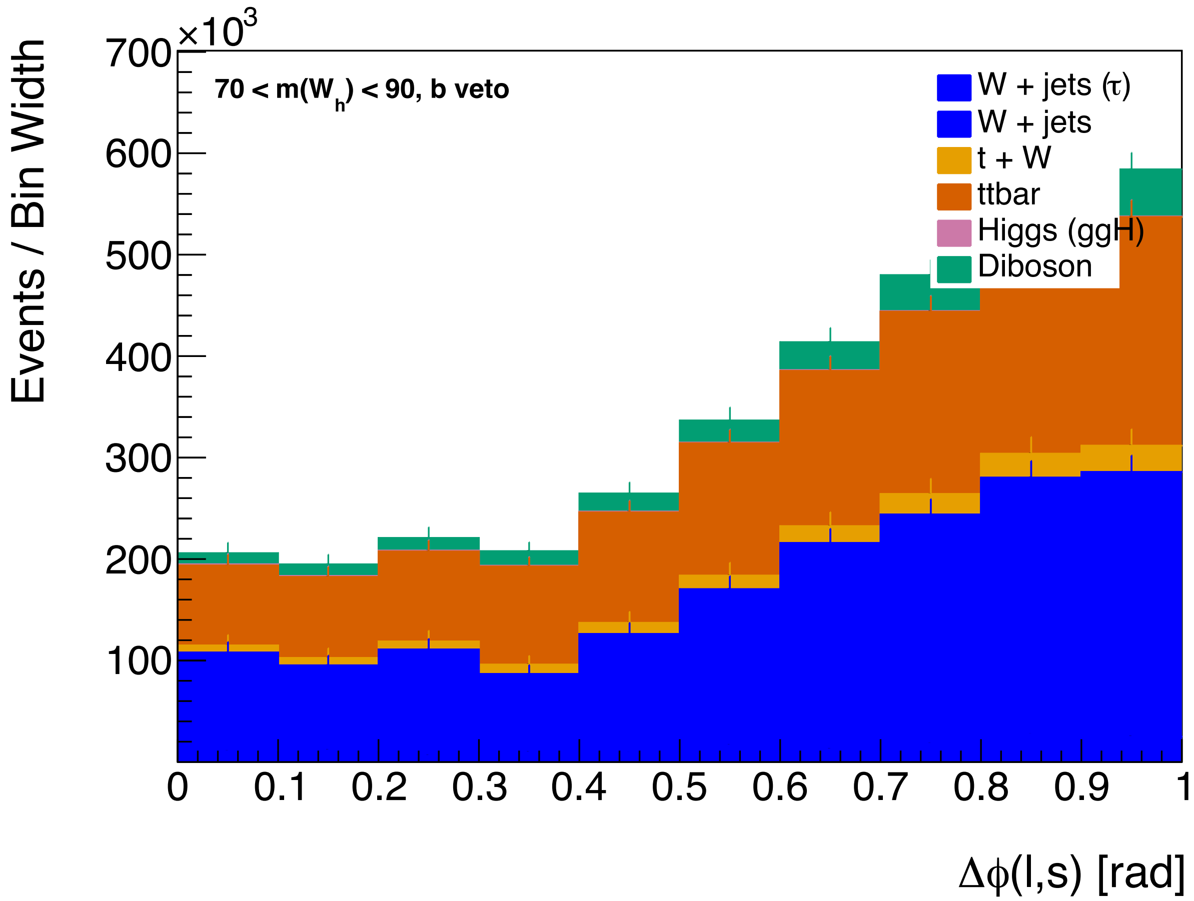


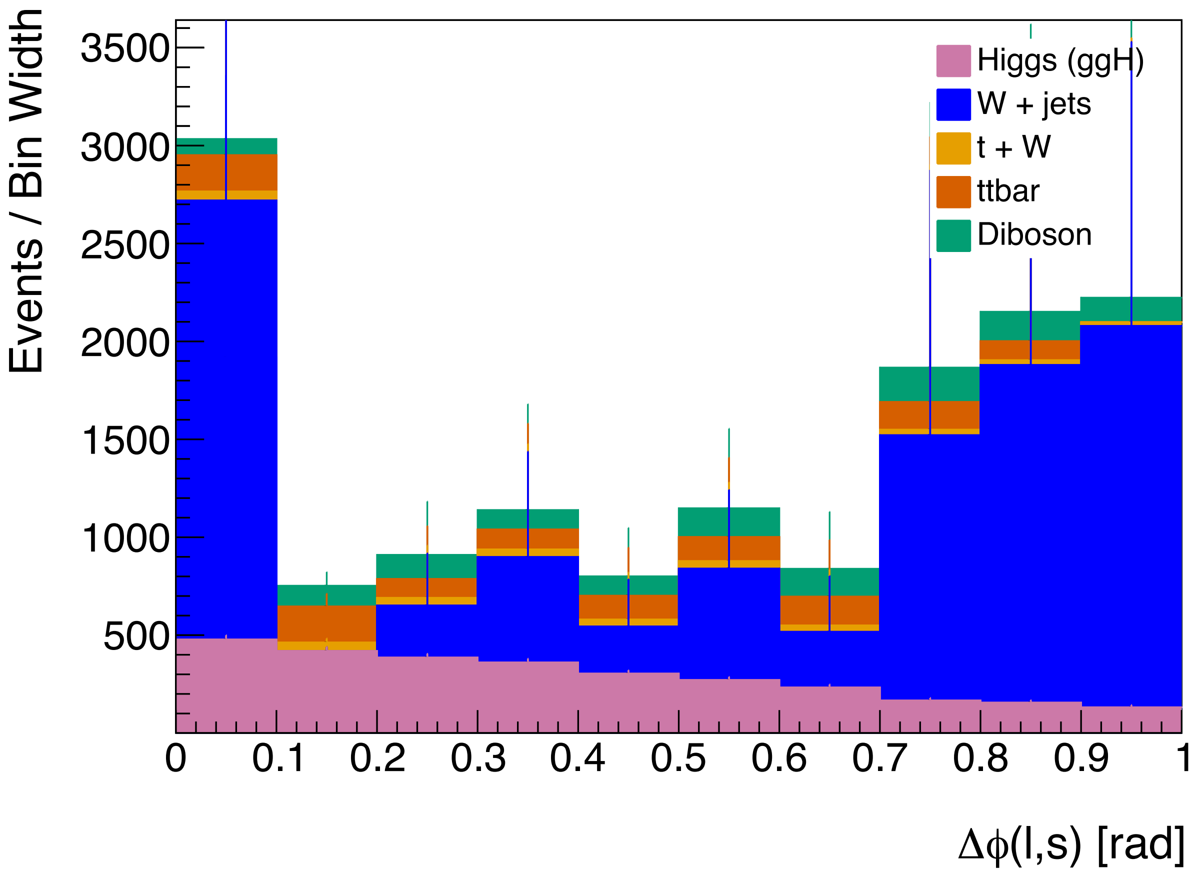


- **Full analyses on real data can optimise these and find additional sensitive selection cuts.**

Observable after cuts

- $\Delta\Phi_{ls}$ before and after applying cuts and selections. The left plot has only on-shell hadronic W-mass and b-veto cuts.This is at 300 fb^{-1}.





- Can see that the backgrounds are significantly suppressed and the signal becomes more dominant.

The W+jets background.

- **With these kinematic & NW cuts, we obtain good S/B** (considering this was, previously, an impossible channel!)

Process	idealised	$\epsilon_c = 40\%$
W + jets	12253 ± 7086	9166 ± 5527
WW	2543 ± 169	1253 ± 118
$t\bar{t}$	723 ± 123	1198 ± 157
tW	213 ± 12	346 ± 15
Higgs	5967 ± 76	2905 ± 53
S/(S+B)	0.28	0.2

- **Can be used for entanglement measurements, but also a brand new Higgs decay topology for free!**

Accessing entanglement

- **Everything here is shamelessly stolen from Alan's paper: https://arxiv.org/pdf/2106.01377.pdf**

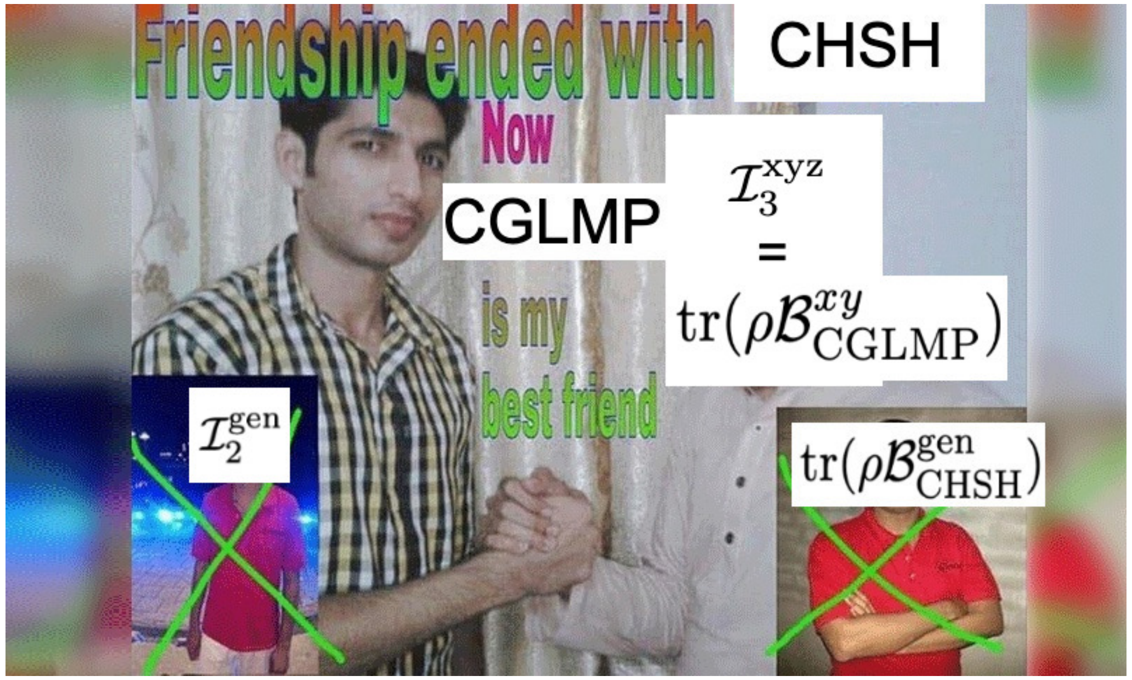
$$\mathrm{tr}(\rho\mathcal{B}^{xy}_{\mathrm{CGLMP}}) = \frac{8}{\sqrt{3}}\left\langle \xi^+_x\xi^-_x + \xi^+_y\xi^-_y \right\rangle_{\mathrm{av}} + 25\left\langle \left((\xi^+_x)^2 - (\xi^+_y)^2\right)\left((\xi^-_x)^2 - (\xi^-_y)^2\right)\right\rangle_{\mathrm{av}} + 100\left\langle \xi^+_x\xi^+_y\xi^-_x\xi^-_y \right\rangle_{\mathrm{av}}$$

$$\xi^{\pm}_i = \hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_{l^{\pm}}$$

$$i = x, y, z$$

$$\mathcal{I}_3^{\mathrm{xyz}} = \max\left(\langle\mathcal{B}^{xy}_{\mathrm{CGLMP}}\rangle, \langle\mathcal{B}^{yz}_{\mathrm{CGLMP}}\rangle, \langle\mathcal{B}^{zx}_{\mathrm{CGLMP}}\rangle\right)$$

- **We have entanglement if $I_3 > 2$.**



Observable Performance

- **Inequality becomes three observables:**

$$\mathrm{tr}(\rho \mathcal{B}^{xy}_{\mathrm{CGLMP}}) = \frac{8}{\sqrt{3}} \left\langle \xi^+_x \xi^-_x + \xi^+_y \xi^-_y \right\rangle_{\mathrm{av}}$$

1st

$$+ 25 \left\langle \left((\xi^+_x)^2 - (\xi^+_y)^2 \right) \left((\xi^-_x)^2 - (\xi^-_y)^2 \right) \right\rangle_{\mathrm{av}}$$

2nd

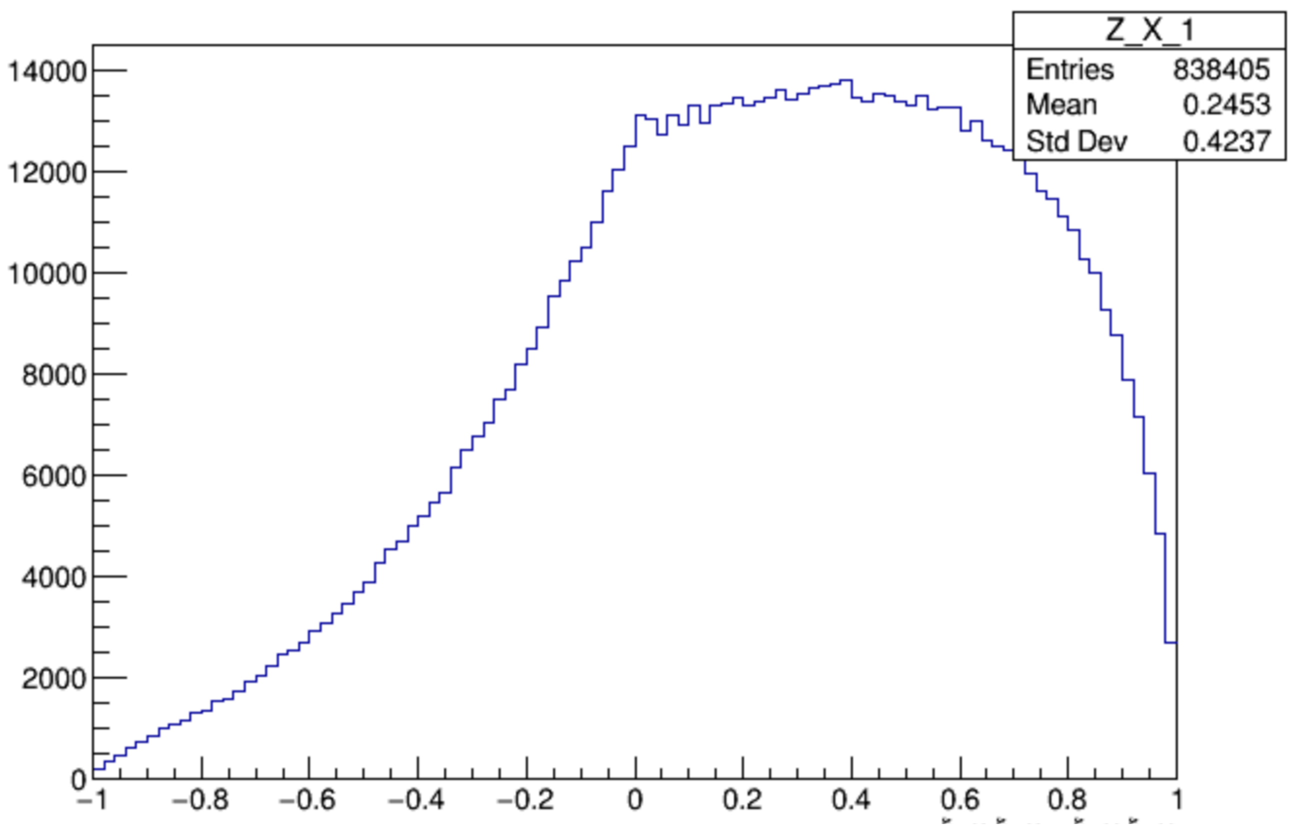
$$+ 100 \left\langle \xi^+_x \xi^+_y \xi^-_x \xi^-_y \right\rangle_{\mathrm{av}}$$

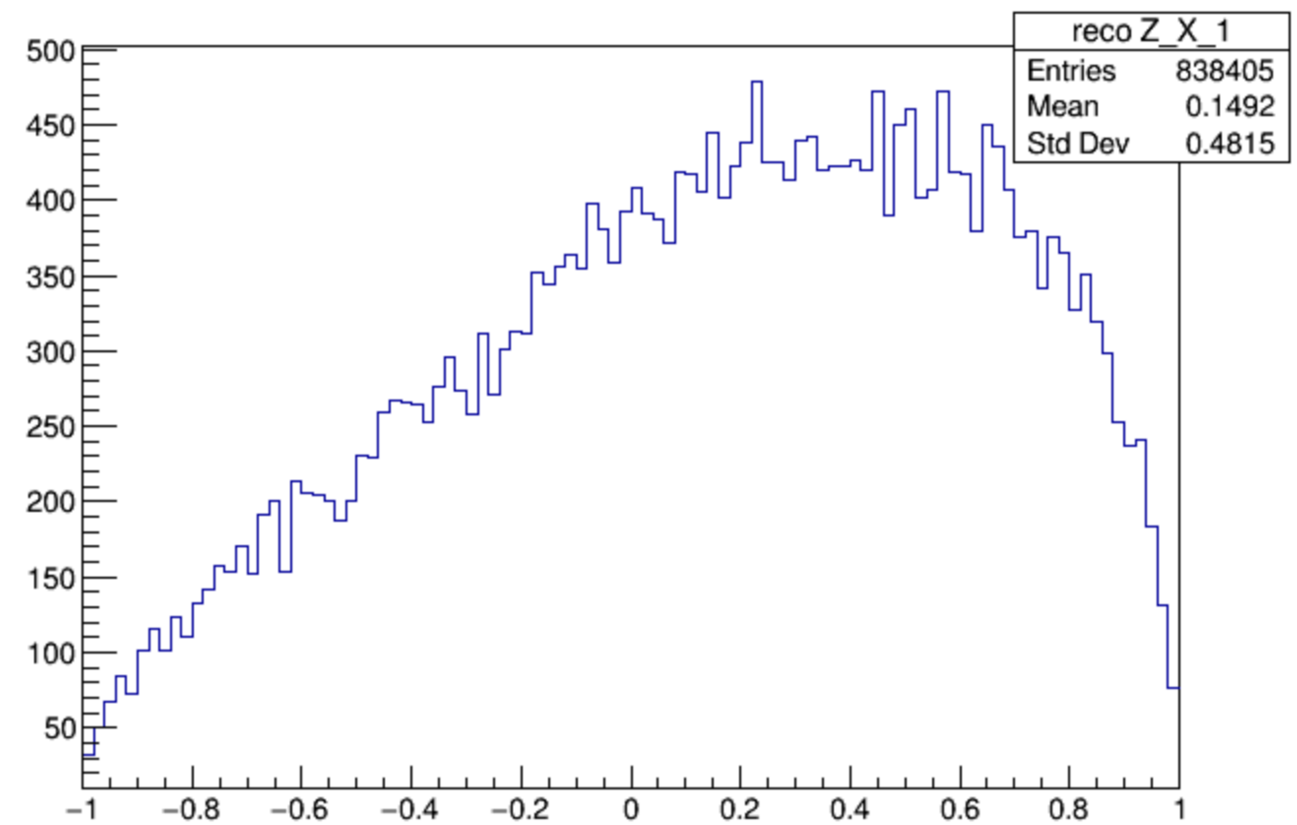
3rd

Observable Performance

First observable

Left is the parton level and right our reco level

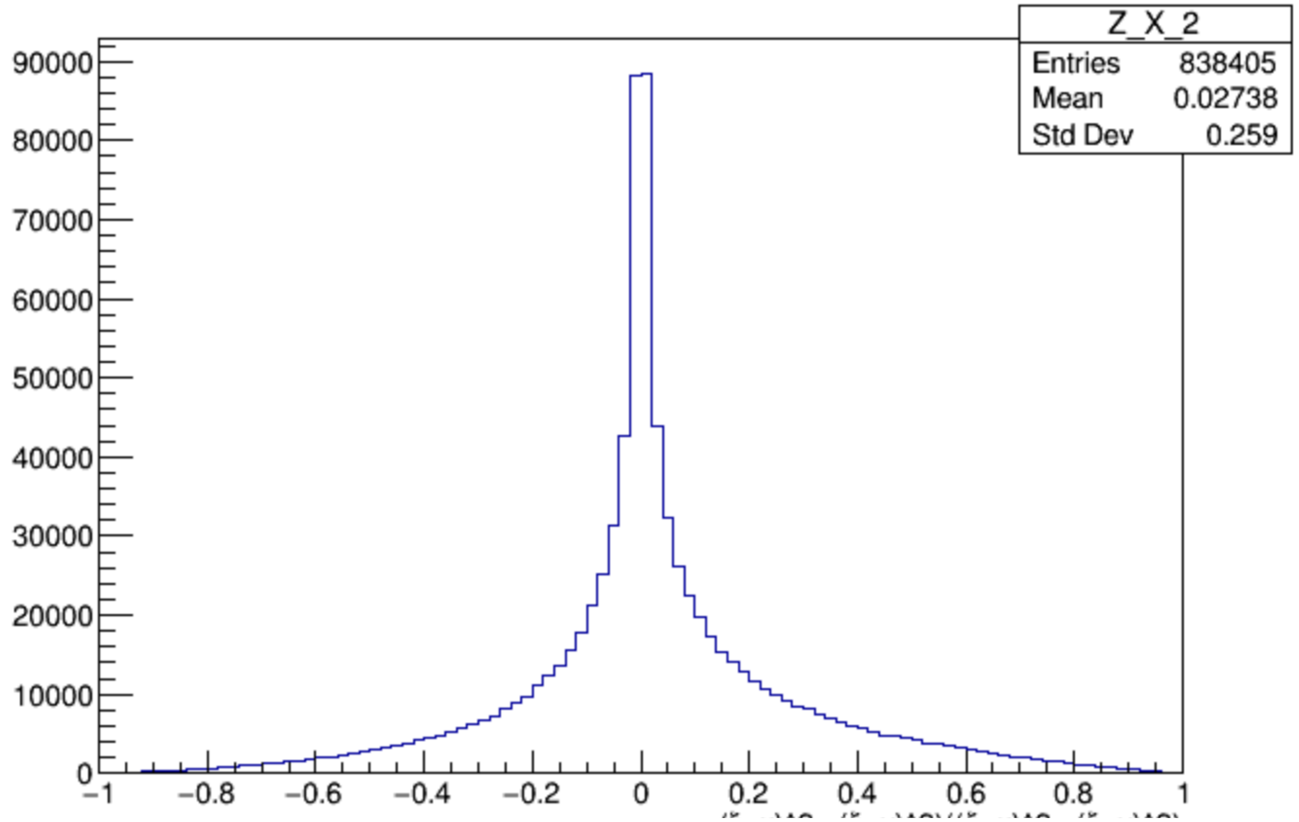


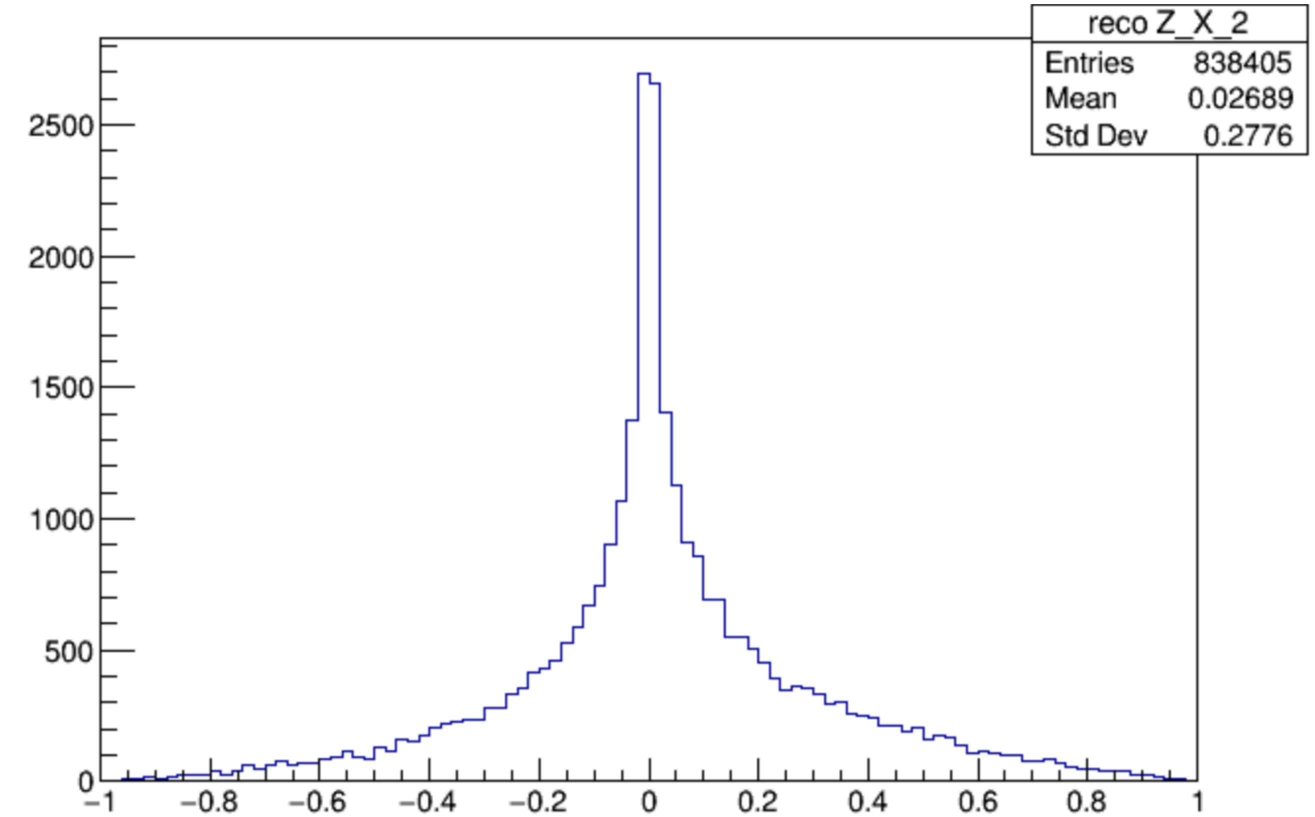


Observable Performance

Second observable

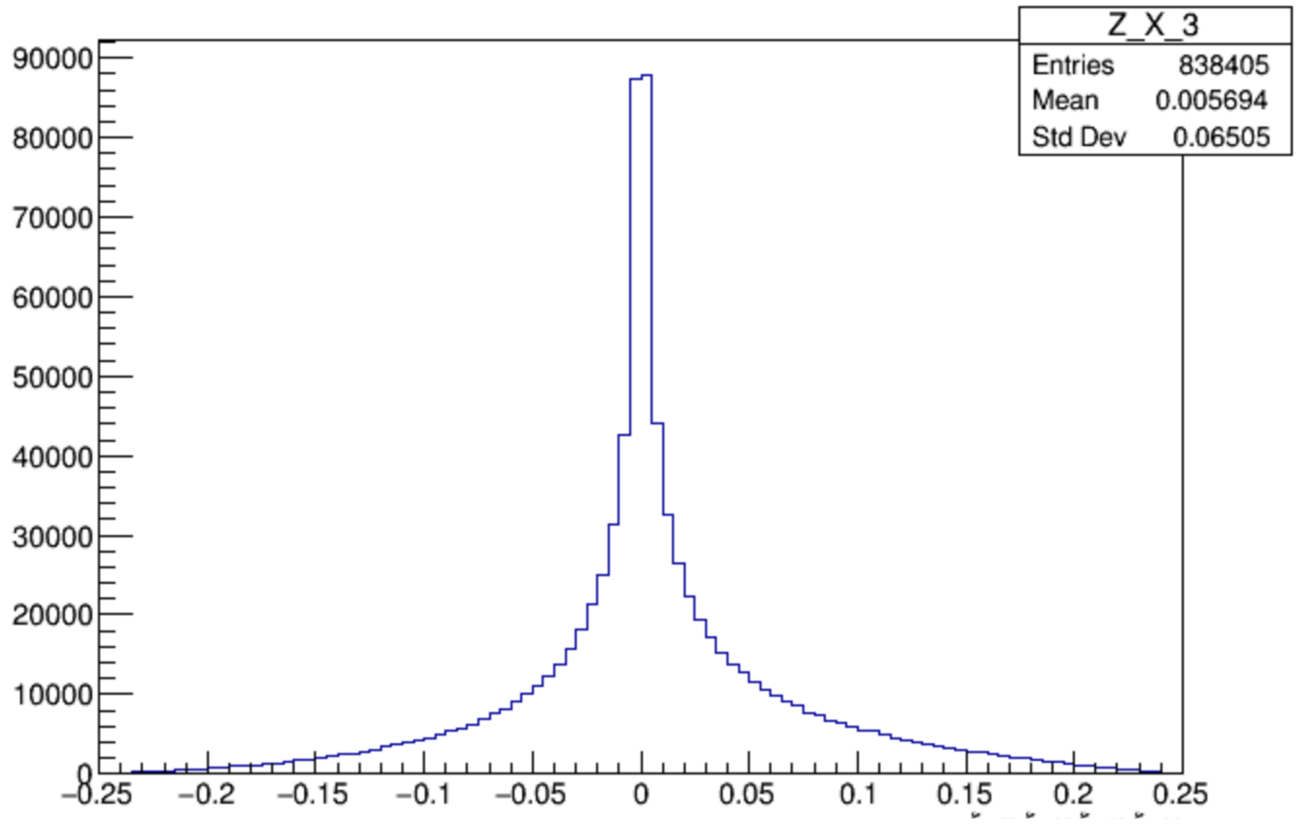
Left is the parton level and right our reco level

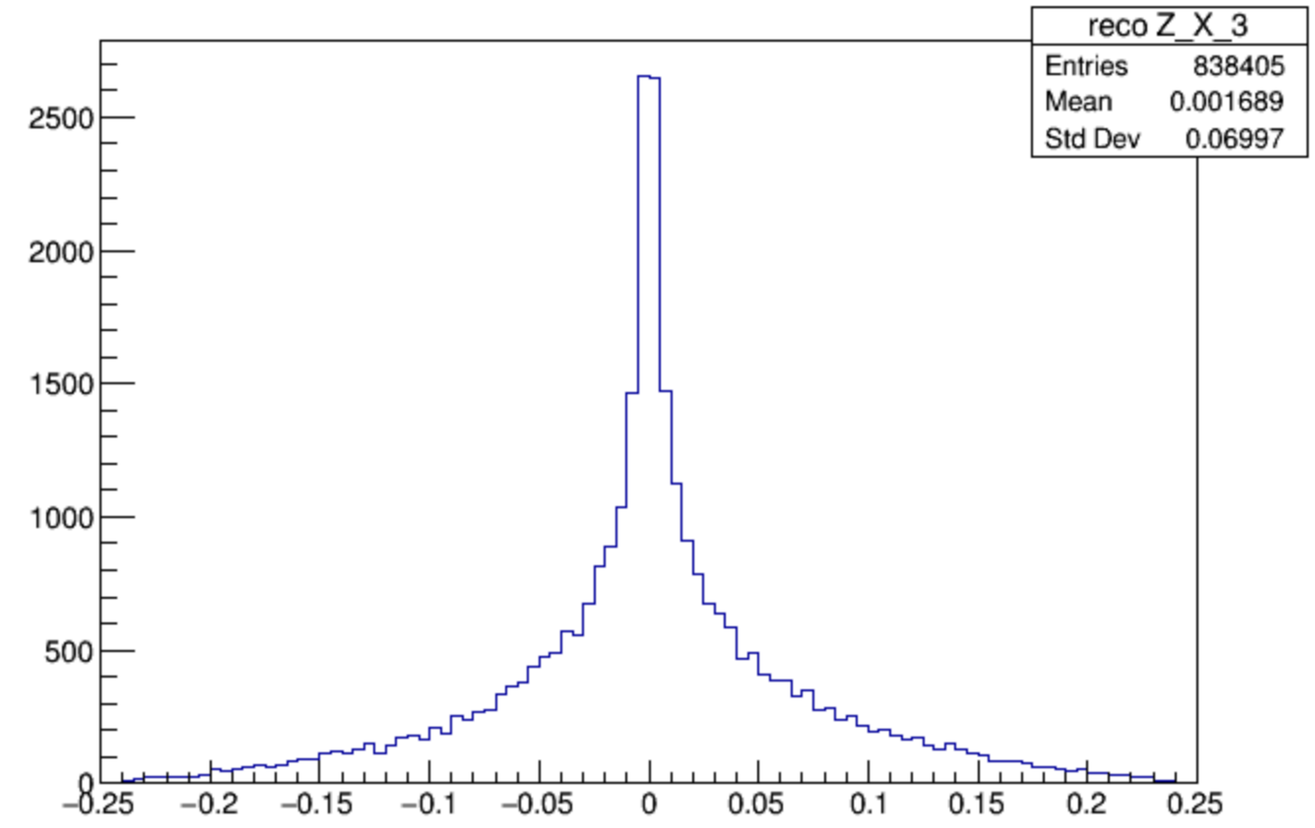




Third observable

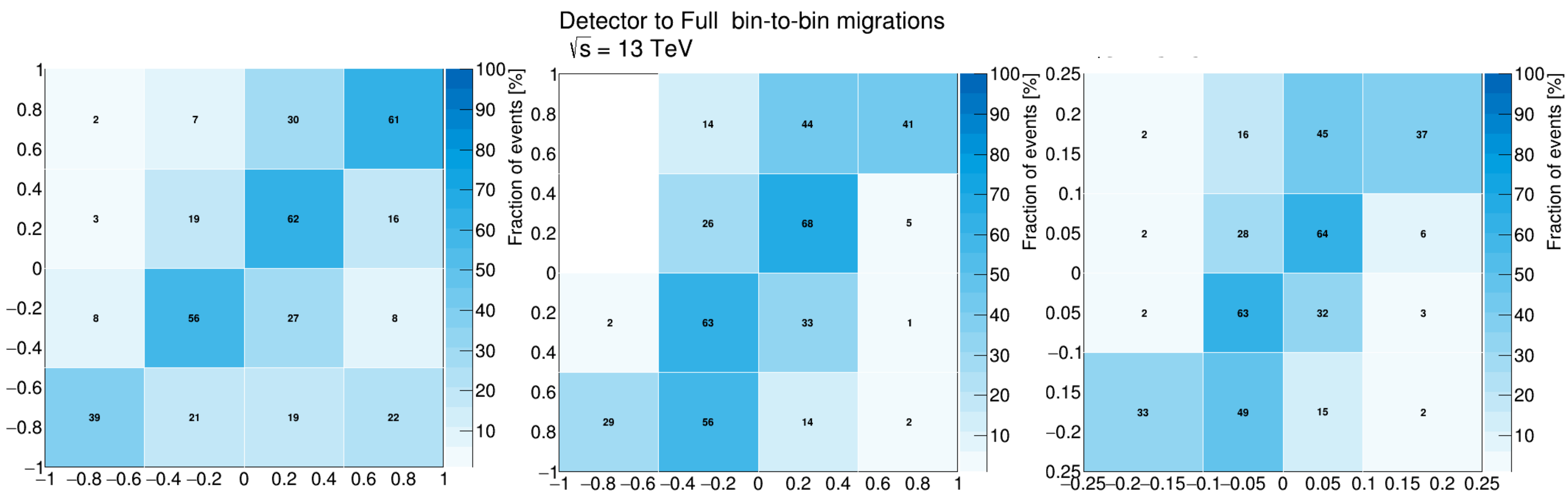
Left is the parton level and right our reco level





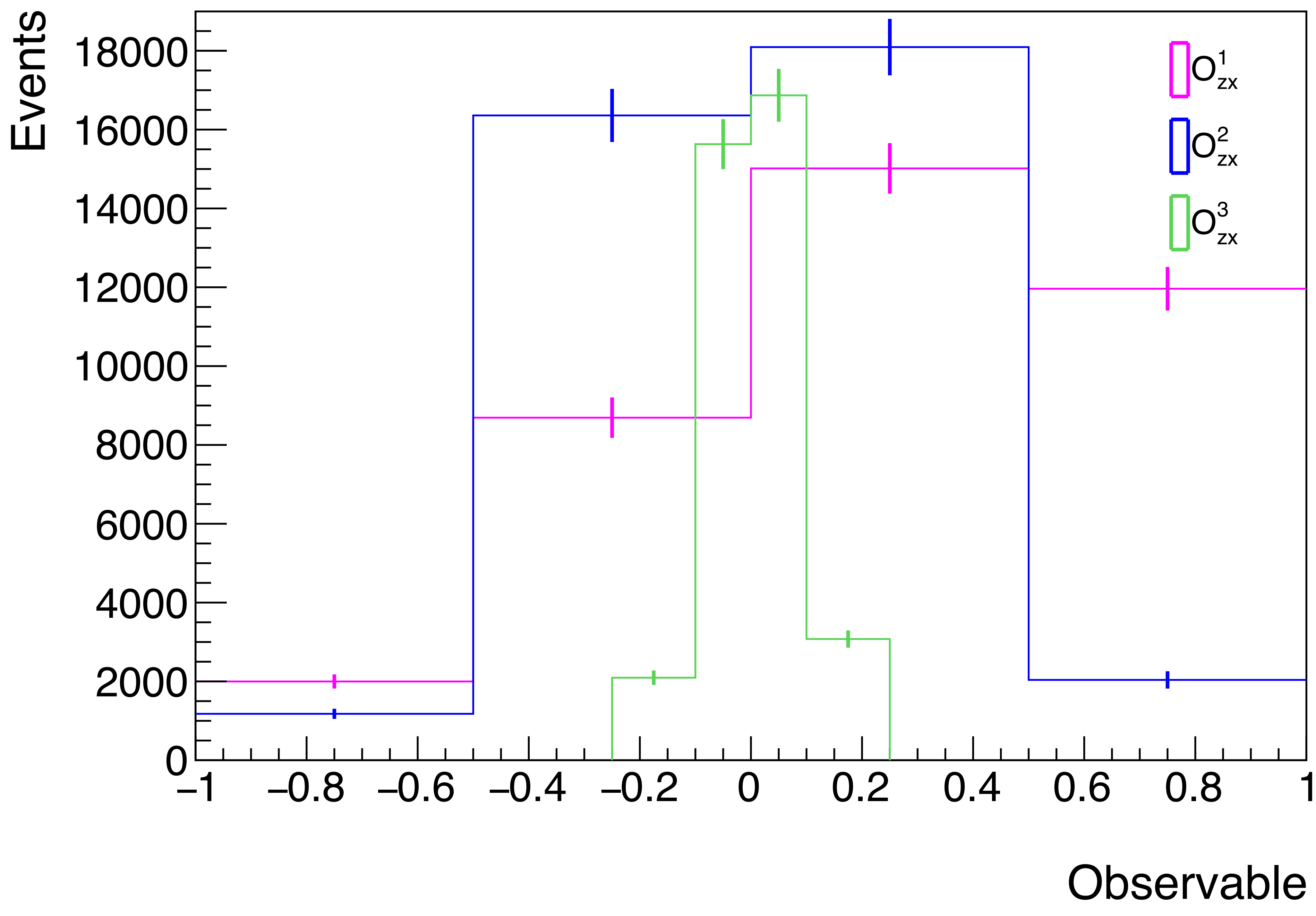
Parameter Extraction

- **Unfolding of the three observables is performed**



Parameter Extraction

- **Unfolding of the three observables is performed**



Parameter Extraction

- Use the unfolded distributions to extract the value for the CGLMP inequality at different integrated luminosities.
- Compute the significance for $I_3 \geq 2$

Luminosity [fb^{-1}]	$\langle \mathcal{B}^{zx}_{\text{CGLMP}} \rangle$ (idealised)	Significance (idealised)
139	2.45 ± 0.25 (0.18)	1.8 (2.5)
300	2.45 ± 0.17 (0.12)	2.65 (3.75)
3000	2.45 ± 0.05 (0.04)	9.0 (11.25)

- Upon introducing jet smearing, the significance for 300 fb^{-1} goes to 2.36.

A real analysis

- The usage of NW can be improved with Machine Learning. Important to note that NW is somewhat idealised in this study.
- Charm tagging can be optimised in an actual analysis. We can expect better S/B ratio and more total signal events.
- Custom trigger could improve acceptance of events with low energy leptons.

Conclusions

- **Potential for an entirely new Higgs decay mode!**
- **Entanglement in $H \rightarrow WW^* \rightarrow jjl\nu$ channel is challenging, but most of these challenges are experimental** (and we're quite good at beating expectations for these types of problems)**.**

Backup

Full selection

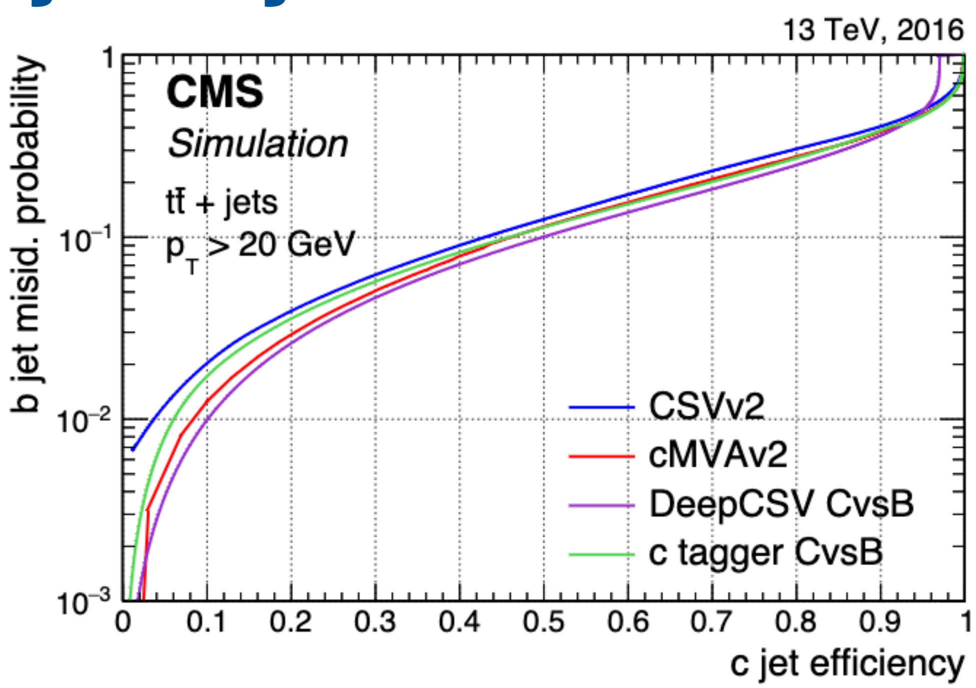
- Pre-selection:
  - Exactly 1 lepton with pT > 20 GeV
  - Exactly 0 b-tagged jets
- C-tagging selection:
  - 2 or more jets, exactly one of which must be c-tagged.
  - At least 1 (c-jet,l-jet) pair with |m_wh - 80.6| < 10 GeV
  - A reconstructed leptonic W boson from NW with weight > 0.7.
  - Maximum two light jets
  - m_cl < 80 GeV (where l is the lepton)

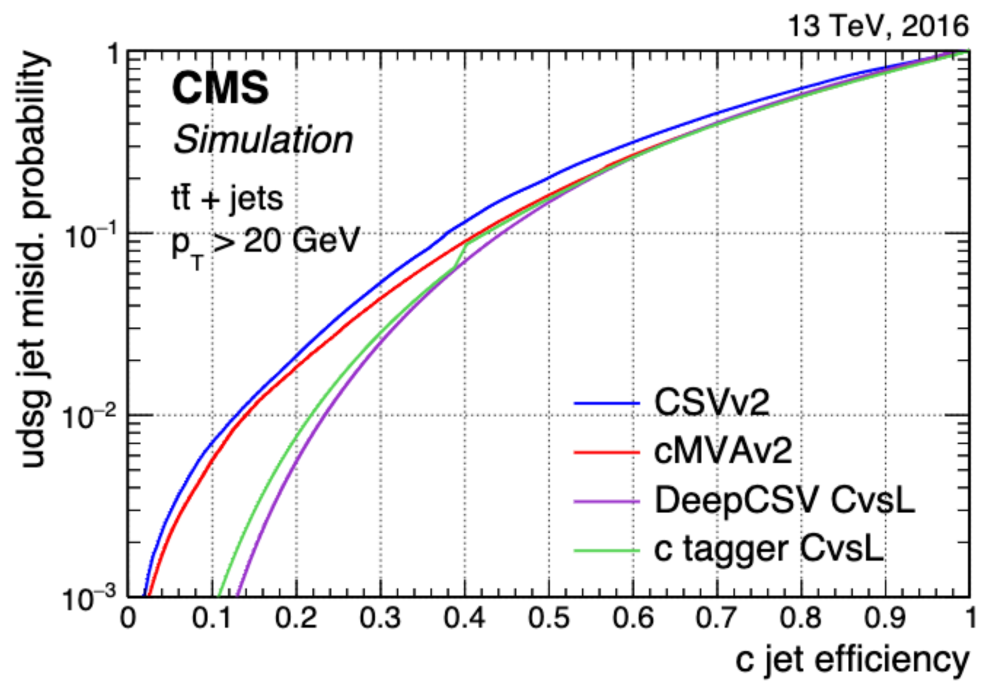
Charm Tagging

- **How well can be tag charm jets at ATLAS and CMS?**

Working point	ε_c (%)	ε_b (%)	ε_{udsg} (%)
c tagger L	88	36	91
c tagger M	40	17	19
c tagger T	19	20	1.2

Suspiciously high numbers that probably come at the cost of high jet rejection.





Taken from https://arxiv.org/pdf/1712.07158.pdf

Backup

- **NW improvements**