Study of polarization fractions in same-sign W boson scattering

Prasham Jain, Beate Heinemann, Oleg Kuprash

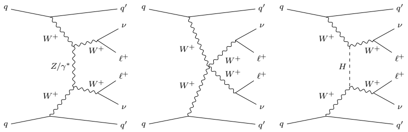
Albert-Ludwigs-Universität Freiburg

GRK Fall Workshop
September 30, 2022

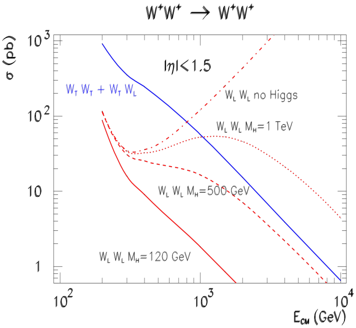
Outline

Introduction: Polarized same-sign $WW(ssWW)$

- Measurement of longitudinally polarized *ssWW* production is sensitive to the way electroweak symmetry is broken
- At high energies, Higgs boson with mass < 1 TeV preserves the unitarity of the tree-level amplitude of $W_L W_L \rightarrow W_L W_L$
- Only purely longitudinal scattering breaks unitarity without Higgs boson
- Measurement of $W_L W_L$ polarization fraction can probe deviations from SM: BSM physics (additional Higgs, new couplings, new resonances)



Representative Feynman diagrams



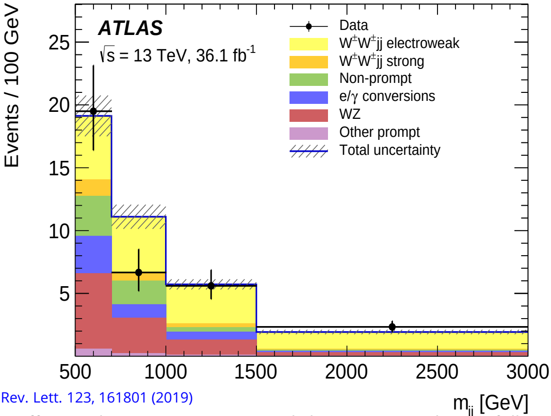
M.Szleper, arXiv:1412.8367

Same-sign WW process: Event Selections

Event Signature: 2 forward jets + 2 same-sign leptons + E_T^{miss}

Event Selections
2 same-sign leptons
$p_T^\ell > 27$ GeV
Veto if ≥ 3 lepton
$m_{\ell\ell} > 20 GeV$
$\lvert m_{\ell\ell} - m_Z \rvert > 15$ GeV
$n_{\text{jets}} \geq 2$
$p_T^{j1(j2)} > 65(35)$ GeV
$E_T^{\text{miss}} > 30$ GeV
$n_{\text{bjets}} = 0$
$\lvert \Delta y_{jj} \rvert > 2$
$m_{jj} \geq 500$ GeV

Observation with 6.5σ using 2015-2016 data



Phys. Rev. Lett. 123, 161801 (2019)

- Differential cross sections currently being extracted using full Run-2 data
- Analysis in review stage, data unblinded

Same-sign WW process: Backgrounds

$W^{\pm}Zjj$ background ($W^{\pm}Zjj \to \ell^{\pm}\nu\ell^{\pm}\ell^{\mp}jj$)

- Dominant background, $\ell^{\mp}$ (from Z decay) out of detector acceptance or not identified
- Estimation $\to$ Sherpa 2.2.2 MC with data-driven Mjj shape correction

Non-prompt background

- Non-prompt/fake lepton: Any object, which is not a prompt lepton, reconstructed as a lepton in the detector
  - Main sources: W+jets and $t\bar{t}$ events
- Estimation $\to$ data-driven techniques: fake factor method

Charge flip background

- e charge misidentification because of incorrect track curvature measurements or wrong e-reconstruction
  - Main sources: high p_T tracks, $e^{\pm} \to e^{\pm}\gamma \to e^{\pm}e^{+}e^{-}$
- Estimation $\to$ data-driven method

Photon conversion background ($V\gamma jj$)

- e channel contributions through γ conversions
- Estimation $\to$ Sherpa 2.2.12 $V\gamma$ MC

Contents

Signal Monte Carlo generation

Electroweak polarized *ssWW* Monte Carlo (MC)

- generated using MadGraph v2.9.5 and Pythia v8.245
- process: $pp \rightarrow W_X^{\pm} W_Y^{\pm} jj, W^{\pm} \rightarrow \ell^{\pm}\nu$ (for $XY = LL, LT, TT$)
  - Here L=longitudinal (helicity=0) and T=transverse (helicity=± 1) polarization of W boson
- reference frames considered (W boson polarization is frame-dependent):
  1. partonic center of mass (pCoM)
  2. WW center of mass (WWCoM)

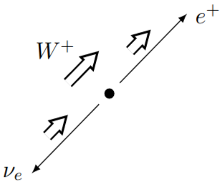
Final States ($+jj$)	Cross-section (fb) pCoM	Cross-section (fb) WWCoM
WW	31.06	
$W_L W_L$	1.98	2.88
$W_T W_L$	10.73	9.37
$W_T W_T$	18.04	18.45
Sum of pol. xsec	30.75	30.70

- Polarization fractions $-f_{LL}, f_{LT}, f_{TT}-$ differ from frame-to-frame

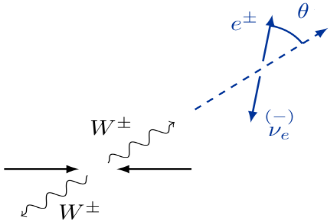
Polarization and its Influence

- Polarization dictates the angular distribution of decay products
  - Used in MC Validation
- W boson only couples to left-handed particles and right-handed anti-particles
  - (W_L, h=0) ℓ^+ escapes $\perp$ W^+ direction
  - ($W_{T,R}$, h=+1) ℓ^+ escapes $\parallel$ W^+ direction
  - ($W_{T,L}$, h=-1) ℓ^+ escapes anti-parallel to W^+ direction

Consider $\boldsymbol{W^+ \rightarrow e^+ \nu_e}$:



Directions of momenta and spins in boson rest frame



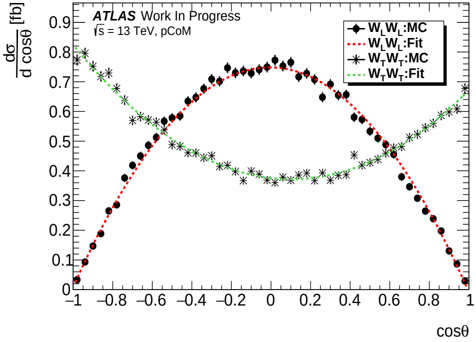
Decay angle of W boson in the W rest frame (denoted in blue)

MC Validation of LL, TT polarization modes

In different lepton flavor events, $\cos\theta$ between the lighter lepton and its parent W is fit to eq. (1).

LL, TT MC Validation fit in $\cos\theta$:

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta} = \frac{3}{8}(1+\cos\theta)^2 f_{T,L} + \frac{3}{8}(1-\cos\theta)^2 f_{T,R} + \frac{3}{4}\sin^2\theta f_L \quad (1)$$



Fit results:

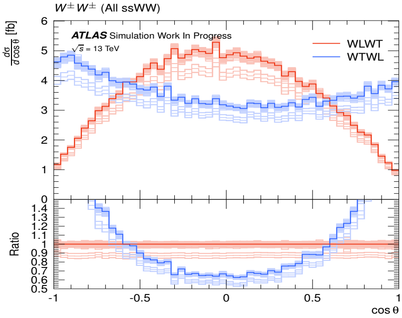
$\sum f_X = 1$ obtained in both frames and

$\implies$ in $W_L W_L$ MC fit $\rightarrow f_L \approx 1$, & in $W_T W_T$ MC fit $\rightarrow (f_{T,L} + f_{T,R}) \approx 1$.

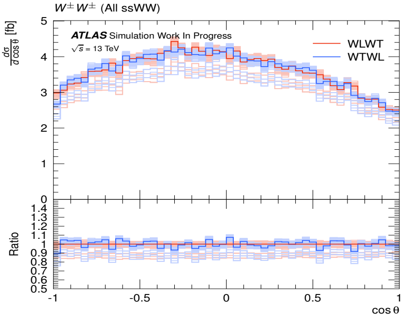
MC Validation of LT polarization mode

MadGraph bug (fixed): [Launchpad Ticket]

- Mixed polarization mode samples$-LT \& TL-$ should be equivalent
- Bug: when final states include τ lepton, polarization preference was given based on syntax
  - for LT syntax, $W_T \to \tau\nu$ always
  - for TL syntax, $W_L \to \tau\nu$ always



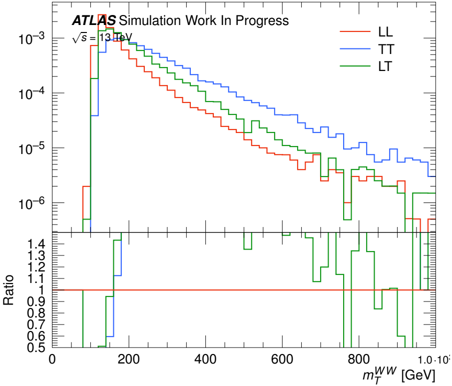
(a) Before fix: polarization preference exists

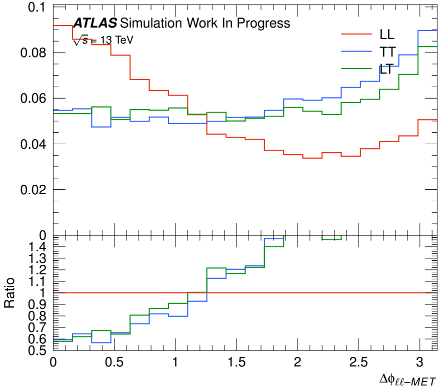


(b) After fix: no polarization preference

Contents

- Important variables include invariant masses, p_T, $\Delta\phi$, ΔR, etc. of final state particles
- Examples of normalized plots (W polarization defined in pCoM frame):





(Complete list of distributions in backup slides)

Contents

DNN for classification of $W_L W_L$ vs $W_X W_T$

Statistics

Label	Class	Pol. modes	SR evts (/100k evts) pCoM	WWCoM
1	Signal	$W_L W_L$	10821	11560
0	Background	$W_L W_T$	10940	10954
0	Background	$W_T W_T$	10726	10857

Weights Signal has $2\times$ weight to account for unbalanced dataset

Input Variables m_T^{WW}, $m_{\ell\ell}$, $\Delta\phi_{jj}$, $\Delta\phi_{\ell\ell}$, $\Delta\phi_{\ell\ell-E_T^{\text{miss}}}$, $p_T^{\ell_1}$, $p_T^{\ell_2}$, $p_T^{j_1}$, $p_T^{j_2}$, $p_T^{\ell\ell}$, $\frac{p_T^{\ell_1} \cdot p_T^{\ell_2}}{p_T^{j_1} \cdot p_T^{j_2}}$, E_T^{miss}, $z_{\ell_1}^*$, $z_{\ell_2}^*$, $\Delta R_{j_1-\ell\ell}$, $\Delta R_{j_2-\ell\ell}$

Preprocessing Standardized normally distributed data

- Gaussian with zero mean and unit variance

Dataset Split (Training,Testing,Validation) = (60,20,20) in %

Avoiding Overtraining Dropout layers, Early stopping, Model checkpoint

DNN Model Details

Model:

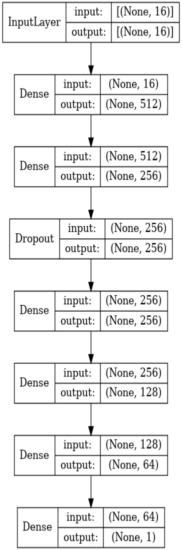
- 5 hidden layers: random normal init., ReLU activation
- 1 dropout layer (40%) to reduce overtraining
- 1 sigmoid output layer with random uniform activation

Compilation:

- Adam Optimizer with learning rate = 0.001
- Loss = binary crossentropy
- Metrics = Accuracy

Model Fit:

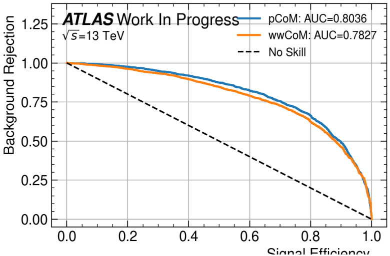
- Batch Size = 1000
- Epochs → stops if val. loss doesn't improve for 5 epochs

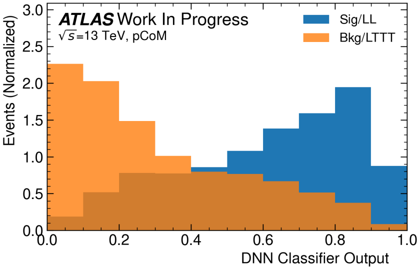


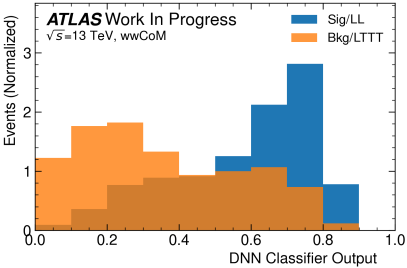
Classification Performance

Trained DNN shows good signal discrimination

- Tested on statistically different sample of SR evts (/100k evts)



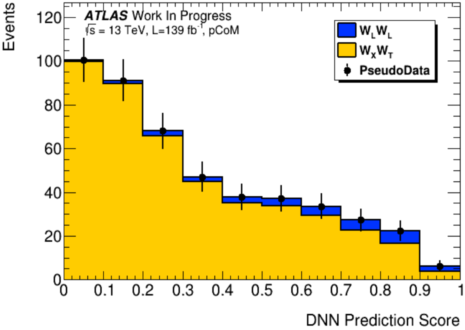




Extracting $W_L W_L$ polarization fraction

χ^2 fit for $W_L W_L$ polarization fraction (Signal Region):

$$\chi^2(c_{LL}) = \sum_{i \in \text{bins}} \frac{\left((c_{LL} \cdot N^i_{LL} + N^i_{XT}) - N^i_{\text{pdata}}\right)^2}{(\Delta N^i_{LL})^2 + (\Delta N^i_{XT})^2 + (\Delta N^i_{\text{pdata}})^2} \quad (2)$$



In SM MC, $c_{LL} = 1$.

$$f_{LL} = \frac{c_{LL} \cdot \sum N_{LL}}{c_{LL} \cdot \sum N_{LL} + \sum N_{XT}} \quad (3)$$

$$f_{LL} = 0.062^{+0.027}_{-0.029} \quad (pCoM)$$

$$f_{LL} = 0.100^{+0.029}_{-0.032} \quad (WWCoM)$$

Note: Pseudodata$= \sum$ all polarization modes

*Other backgrounds, systematics, etc. are yet to be included.

Contents

Classification: $W_L W_L$ vs $W_X W_T$+Other Bkgs

Dataset DNN trained with reco-level information

Statistics

Label	Class	No. of events pCoM	No. of events WWCoM
1	Signal	19673	21643
0	Background	131343	132210

Signal LL pol. ($W_L^{\pm} W_L^{\pm}$)

Background TL, TT pol., $W^{\pm}W^{\pm}$ QCD, WZ QCD, WZ EW6, charge Flip

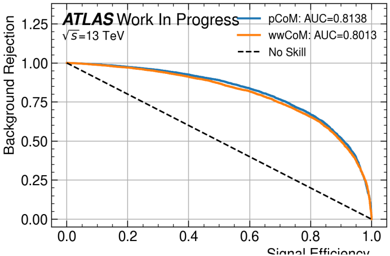
Weights class_weight(sig) = n(Bkg)/n(Sig) for unbalanced dataset

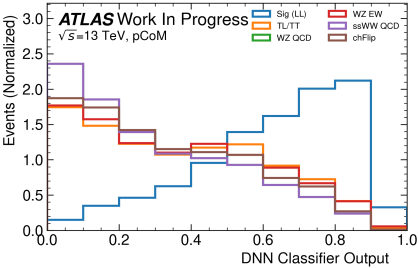
Misc. Features Input variables, preprocessing, dataset split, overtraining checks, etc. are unchanged

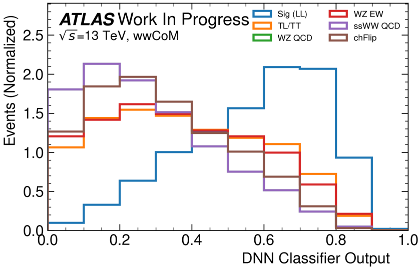
Classification Performance

Trained DNN shows good signal discrimination

- Tested on the entire dataset of SR evts and backgrounds



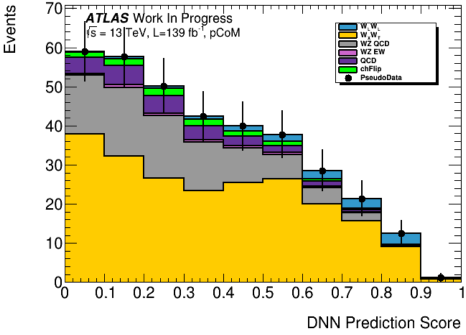




Extracting $W_L W_L$ polarization fraction

χ^2 fit for $W_L W_L$ polarization fraction (Signal Region):

$$\chi^2(c_{LL}) = \sum_{i \in \text{bins}} \frac{\left((c_{LL} \cdot N^i_{LL} + N^i_{\text{Bkgs}}) - N^i_{\text{pdata}}\right)^2}{(\Delta N^i_{LL})^2 + (\Delta N^i_{\text{Bkgs}})^2 + (\Delta N^i_{\text{pdata}})^2} \quad (4)$$



In SM MC, $c_{LL} = 1$.

$$f_{LL} = \frac{c_{LL} \cdot \sum N_{LL}}{c_{LL} \cdot \sum N_{LL} + \sum N_{XT}} \quad (5)$$

$$f_{LL} = 0.056^{+0.043}_{-0.048} \quad (pCoM)$$
$$f_{LL} = 0.089^{+0.049}_{-0.055} \quad (WWCoM)$$

Note: Pseudodata$= \sum$ all backgrounds

*Systematics and some other backgrounds yet to be included.

Summary

Electroweak polarized *ssWWjj* production:

- MC Samples successfully generated and validated

Classification of $W_L W_L$ polarization modes:

- Deep Neural Networks constructed to classify
  - $W_L W_L$ vs $W_X W_T$ at truth-level
  - $W_L W_L$ vs Bkgs at reco-level
- $W_L W_L$ polarization fraction:

Frame	Fraction	Truth-level Result	Reco-level Result
*p*CoM	f_{LL}	$0.062^{+0.027}_{-0.029}$	$0.056^{+0.043}_{-0.048}$
*WW*CoM	f_{LL}	$0.100^{+0.029}_{-0.032}$	$0.089^{+0.049}_{-0.055}$

- Outlook:
  - Include remaining backgrounds and systematics
  - Optimize the classification algorithm further

Thank you for your attention!

Backup slides

MC generation: MadGraph

MadGraph commands for MC generation:

```
generate p p > w+{X} w+{Y} j j QED=4 QCD=0, w+ > l+ vl @1
add process p p > w+{X} w+{Y} j j QED=4 QCD=0, w+ > l+ vl, w+ > ta+ vl @1
add process p p > w+{X} w+{Y} j j QED=4 QCD=0, w+ > ta+ vl, w+ > l+ vl @2
add process p p > w+{X} w+{Y} j j QED=4 QCD=0, w+ > ta+ vl @1
add process p p > w-{X} w-{Y} j j QED=4 QCD=0, w- > l- vl~ @1
add process p p > w-{X} w-{Y} j j QED=4 QCD=0, w- > l- vl~, w- > ta - vl ~ @1
add process p p > w-{X} w-{Y} j j QED=4 QCD=0, w- > ta- vl~, w- > l - vl ~ @2
add process p p > w-{X} w-{Y} j j QED=4 QCD=0, w- > ta- vl~ @1
```

$\implies$ using explicit τ decays

MC Cross-sections: Locally generated samples

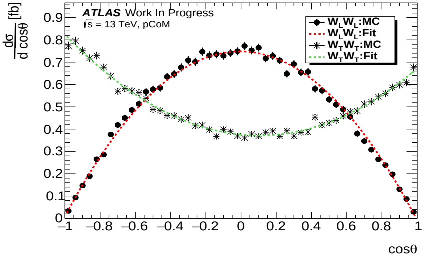
Total events = 100,000 each

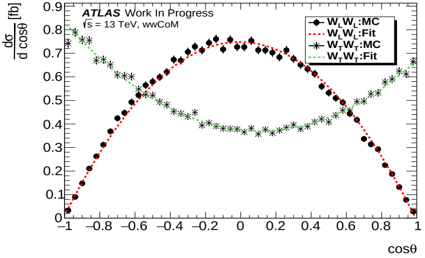
Final States $(+jj)$	**BW cutoff**	**Cross-section (fb)** pCoM	WW CoM
$\ell\nu\ell\nu$	1000	35.17	
WW	15	31.01 ± 0.16	
$W_L W_L$	15	1.97 ± 0.01	2.89 ± 0.02
$W_T W_L$	15	10.78 ± 0.06	9.41 ± 0.05
$W_T W_T$	15	18.06 ± 0.10	18.49 ± 0.10
Sum of pol. xsec		30.80 ± 0.12	30.79 ± 0.11

MC: $\cos\theta$ Validation

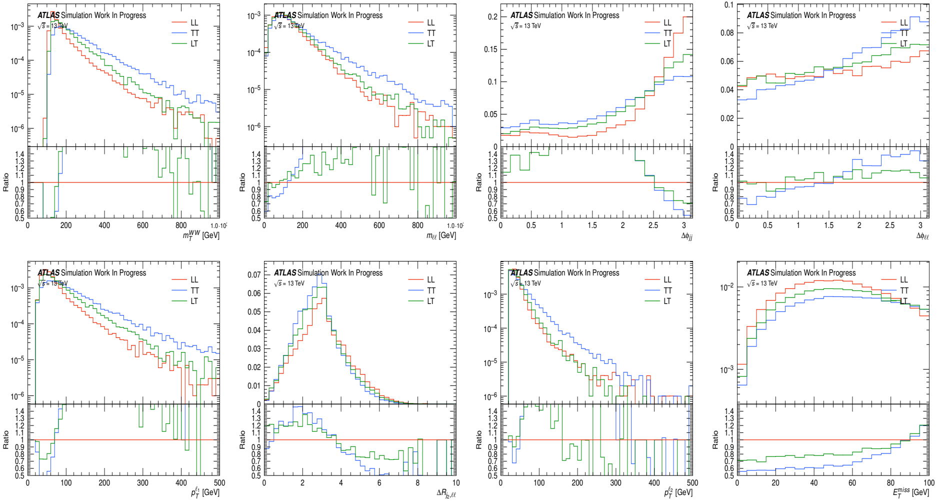
$\cos\theta$ validation fit:

- Polar angle $\theta =$ angle between the flight direction of one of the Ws (rest frame in which sample is generated) and the lepton ℓ it decays into (W's rest frame)
- Opposite flavor leptons, W decaying to lightest ℓ selected
- W, ℓ boosted from lab frame into the rest frame in which sample is generated
- ℓ further boosted to its W's rest frame





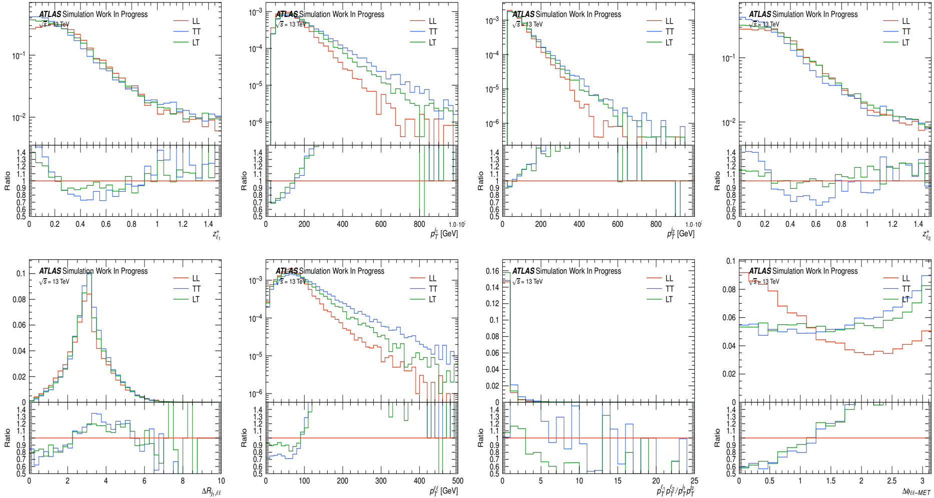
m_T^{WW}, $m_{\ell\ell}$, $\Delta\phi_{jj}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell_1}$, $\Delta R_{j_2-\ell\ell}$, $p_T^{\ell_2}$, E_T^{miss}



Polarization Discriminating Variables

Normalized, p-CoM

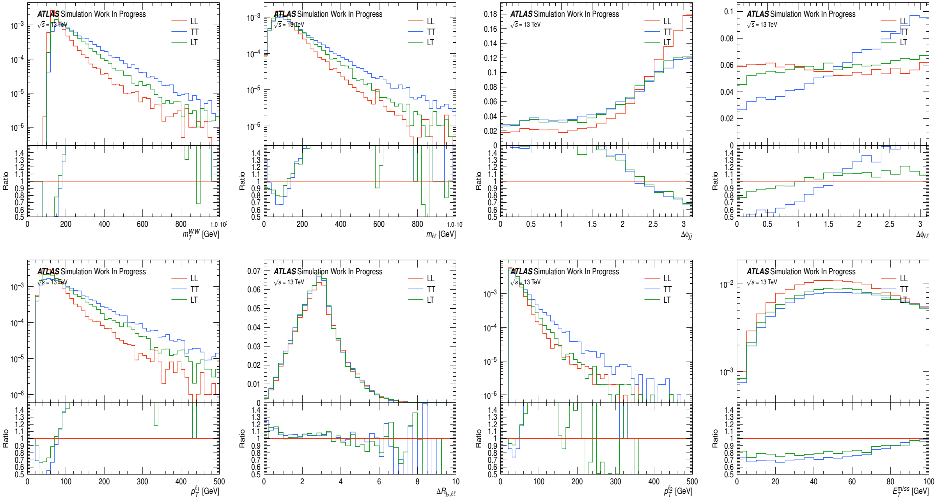
$z^*_{\ell_1}$, $p_T^{j_1}$, $p_T^{j_2}$, $z^*_{\ell_2}$, $\Delta R_{j_1-\ell\ell}$, $p_T^{\ell\ell}$, $p_T\left(\frac{\ell_1\ell_2}{j_1j_2}\right)$, $\Delta\phi_{\ell\ell-E_T^{\text{miss}}}$



Polarization Discriminating Variables

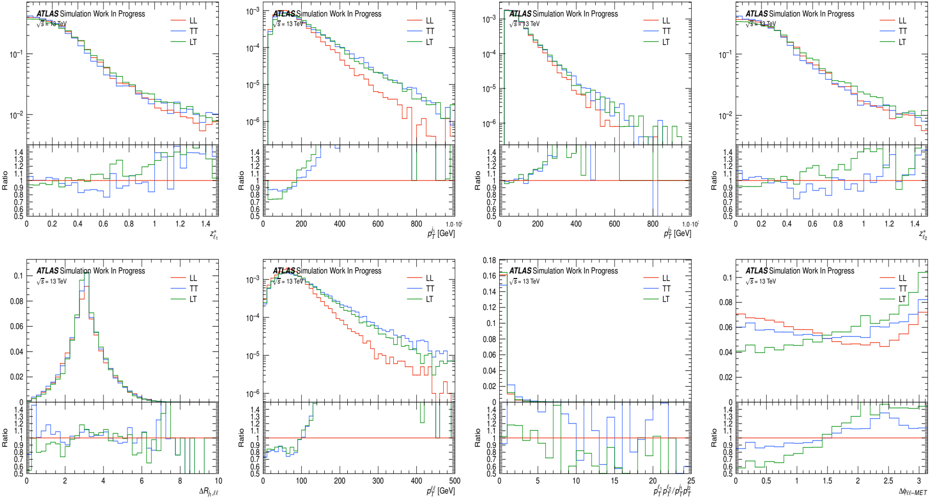
Normalized, WW-CoM

m_T^{WW}, $m_{\ell\ell}$, $\Delta\phi_{jj}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell_1}$, $\Delta R_{j_2-\ell\ell}$, $p_T^{\ell_2}$, E_T^{miss}



Polarization Discriminating Variables   Normalized, WW-CoM

$z^*_{\ell_1}$, $p_T^{j_1}$, $p_T^{j_2}$, $z^*_{\ell_2}$, $\Delta R_{j_1-\ell\ell}$, $p_T^{\ell\ell}$, $p_T(\frac{\ell_1\ell_2}{j_1j_2})$, $\Delta\phi_{\ell\ell-E_T^{\text{miss}}}$



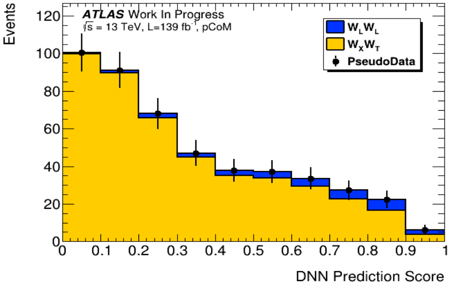
Polarization Discriminating Variables

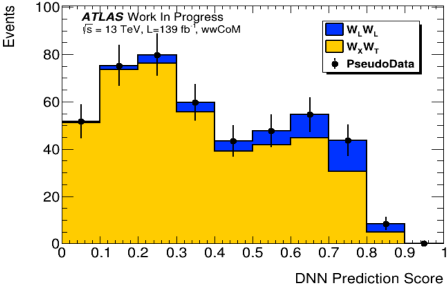
$$m_T^{WW} = \sqrt{(\sum_i E_i)^2 + (\sum_i p_{z,i})^2} \tag{6}$$

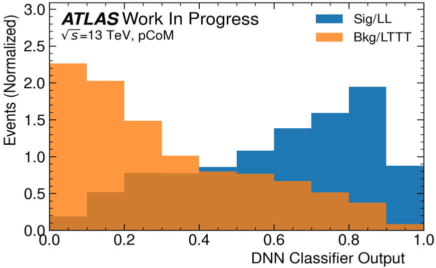
$$z^*_{\ell_1} = \left| \frac{\eta_{\ell_1} - 0.5 \cdot (\eta_{j_1} + \eta_{j_2})}{\Delta \eta_{jj}} \right| \tag{7}$$

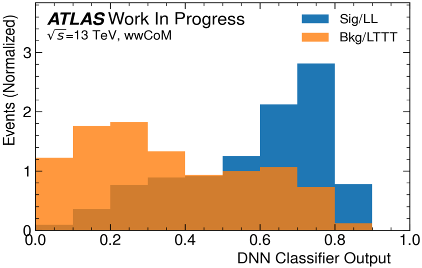
DNN Discriminant (truth-level)

p-CoM, WW-CoM

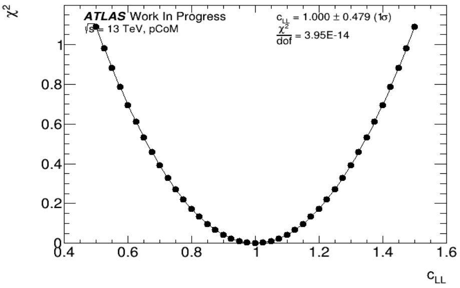


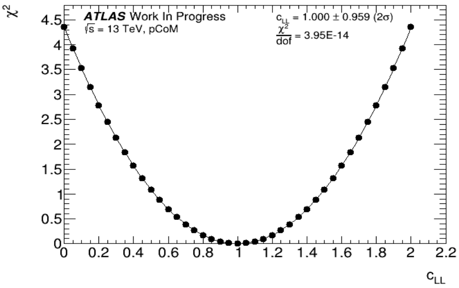


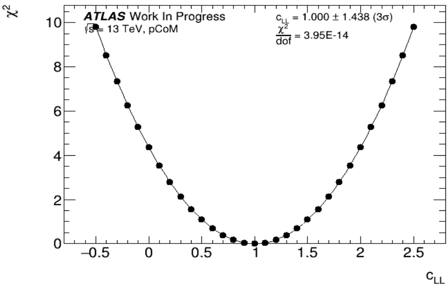




Chi-squared fit (truth-level)

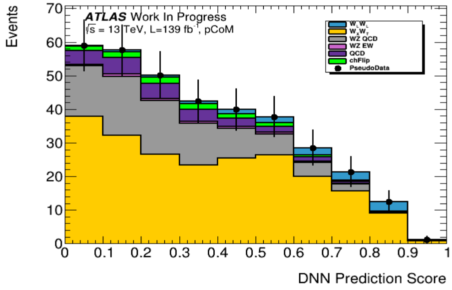


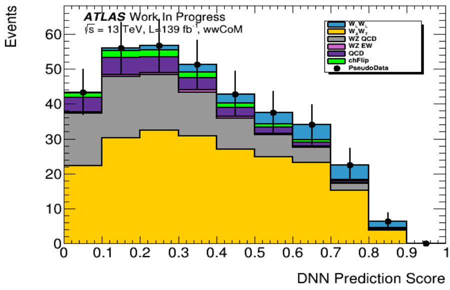


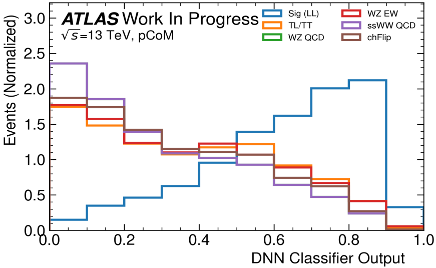


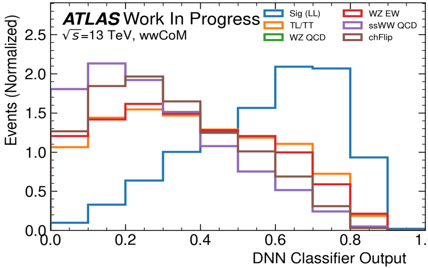
DNN Discriminant (reco-level)

p-CoM, WW-CoM









Chi-squared fit (reco-level)