Study of polarization fractions in same-sign W boson scattering at $\sqrt{s} = 13\,TeV$ with the ATLAS detector

Prasham Jain[1,2]

[1] Deutsches Elektronen-Synchrotron

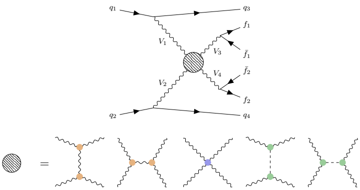
[2] University of Freiburg

DESY SM Meeting
April 22, 2025

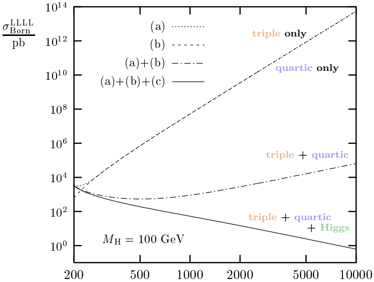
Introduction: Electroweak-induced same-sign $WWjj$ production

Vector boson scattering (VBS) $V_1 V_2 \to V_3 V_4$ process in $W^\pm W^\pm jj$ final states:

- Rare process: low cross-section even at energy scales achieved by the LHC
- $W^\pm W^\pm jj$ characterized by large EWK-to-QCD production mode ratio
- The Higgs mechanism generates gauge boson masses and their longitudinal polarization
- In $V_L V_L \to V_L V_L$, unitarity is violated without Higgs interaction $\implies$ indirect test of electroweak symmetry breaking mechanism
- Allows test of anomalous quartic gauge couplings (aQGCs)
- Probes deviations from the SM: additional Higgs, new couplings, new resonances



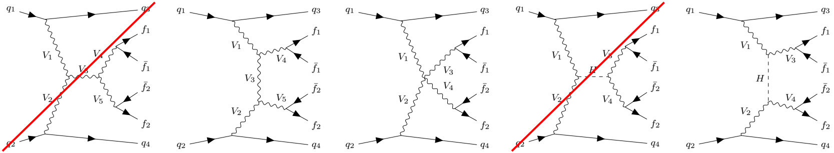
(a) triple (b) quartic (c) Higgs



A. Denner T. Hahn, Nucl.Phys.B525:27-50,1998

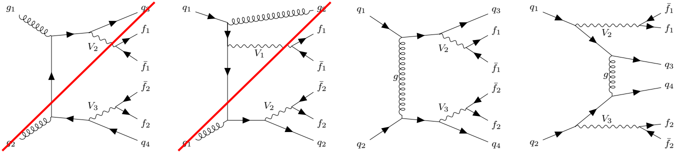
$W^{\pm}W^{\pm}jj$ has the largest EWK-to-QCD production cross-section ratio amongst all the VBS sensitive $VVjj$ final states

- supressed diagrams in $W^{\pm}W^{\pm}jj$ crossed with red line



(a) EW VBS $VVjj$ (b) EW VBS $VVjj$ (c) EW VBS $VVjj$ (d) EW VBS $VVjj$ (e) EW VBS $VVjj$

(f) EW non-VBS $VVjj$ (g) EW non-VBS $VVjj$



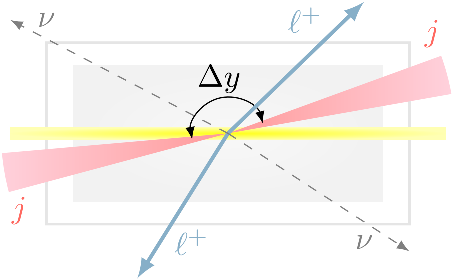
(h) QCD $VVjj$ (i) QCD $VVjj$ (j) QCD $VVjj$ (k) QCD $VVjj$

Same-sign WW process: Event Signature

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

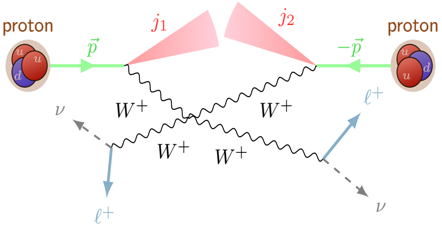
VBS $W^{\pm}W^{\pm}jj$ event topology:

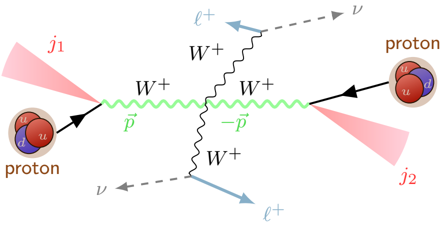
- 2 incoming quarks
- 2 tagging jets (j) in the forward regions with large rapidity gap (Δy)
- 2 outgoing leptons (ℓ) and neutrinos (ν) which come from W boson decays
- Ws lie in the central region
- No additional jet in the gap region



Measurement of *L*-polarized $W^\pm W^\pm jj$ production:

- Polarization is not Lorentz invariant $\implies$ reference frame must be chosen
- W polarization defined in following reference frames:
  - parton-parton center-of-mass (*pCoM*)
  - diboson or *WW* center-of-mass (*WWCoM*)





pCoM **WWCoM**

Polarization and its Influence

Polarization of gauge boson affects the kinematic & angular distributions of decay products

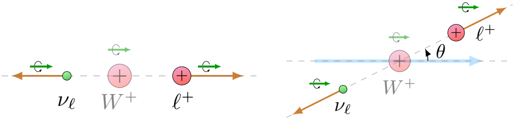
- Used in MC Validation

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 $||$ W^+ direction
- ($W_{T,L}$, h=-1) ℓ^+ escapes anti-parallel to W^+ direction

$$\boxed{\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta} = \underbrace{\frac{3}{8}(1 \mp \cos\theta)^2 f_{T,L} + \frac{3}{8}(1 \pm \cos\theta)^2 f_{T,R}}_{\text{Transverse}} + \underbrace{\frac{3}{4}\sin^2\theta f_L}_{\text{Longitudinal}}, \quad \text{for } W^\pm}$$

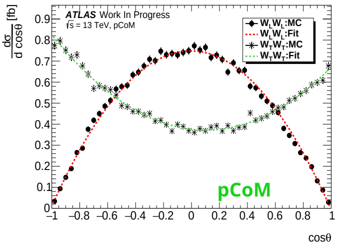
Consider the decay $\boldsymbol{W^+ \to \ell^+ \nu}$:



(a) Helicity in $W^+ \to \ell^+ \nu$ decay in boson rest frame

(b) Polar angle θ

Distribution for W decaying into lighter ℓ



Signal Monte Carlo: Generation & Validation

- *ssWW* EWK MC samples generated using Madgraph ($pp \to W_X W_Y jj$, $W \to \ell\nu$)
  - Wrong polarization assignment for final states with τ leptons
- $\cos\theta$ plots of mixed pol. samples show pol. preference in τ channels
  - $W_L W_T$ sample prefers $W_T \to \tau\nu$
  - $W_T W_L$ sample prefers $W_L \to \tau\nu$
- Contacted MG5 developers [Launchpad Ticket]
  - Received 2 possible solutions: (1) Syntax change (implemented), (2) Patch
- Generated $W_L W_T$, $W_T W_L$ samples with suggested syntax correction
  - Old syntax `(l=e,mu,tau)`

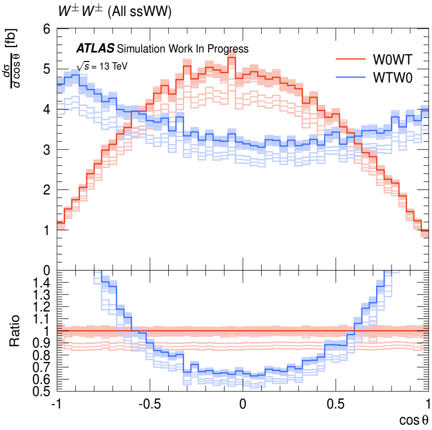
```
generate p p > w+{X} w+{Y} j j QED=4 QCD=0, w+ > l+ vl @1
add process p p > w-{X} w-{Y} j j QED=4 QCD=0, w- > l- vl~ @1
```

  - New syntax `(l=e,mu)`

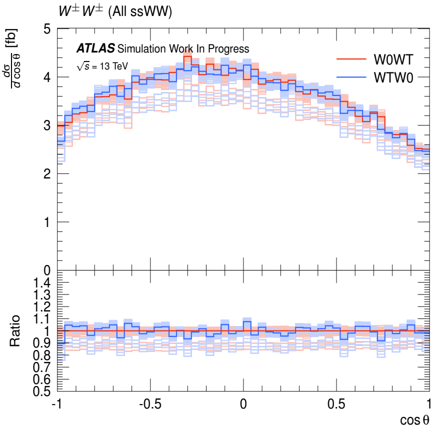
```
generate p p > w+{0} w+{T} j j QED=4 QCD=0, w+ > l+ vl @1
add process p p > w+{0} w+{T} j j QED=4 QCD=0, w+ > l+ vl, w+ > ta+ vl @1
add process p p > w+{0} w+{T} j j QED=4 QCD=0, w+ > ta+ vl, w+ > l+ vl @2
add process p p > w+{0} w+{T} j j QED=4 QCD=0, w+ > ta+ vl @1
add process p p > w-{0} w-{T}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

Comparison of new and old $\cos\theta$ plots: All ssWW events

Note: $W_0 = W_L$



(a) Before Fix



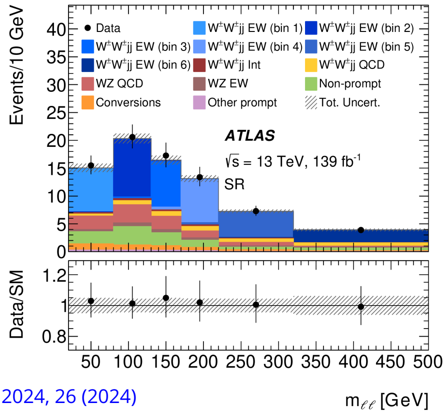
(b) After Fix

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
$n_{\mathrm{jets}} \geq 2$
$p_T^{j1(j2)} > 65(35)$ GeV
$E_T^{\mathrm{miss}} > 30$ GeV
$n_{\mathrm{bjets}} = 0$
$\lvert \Delta y_{jj} \rvert > 2$
$m_{jj} \geq 500$ GeV

VBS enhancing selections

Current ATLAS result:

EWK $ssWW$ cross section measurement using full Run 2 data



JHEP 2024, 26 (2024)

$W^{\pm}Zjj$ 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

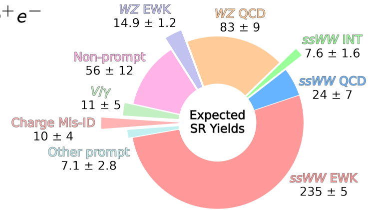
Charge flip/misidentification background

- 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 ($V\gamma jj$)

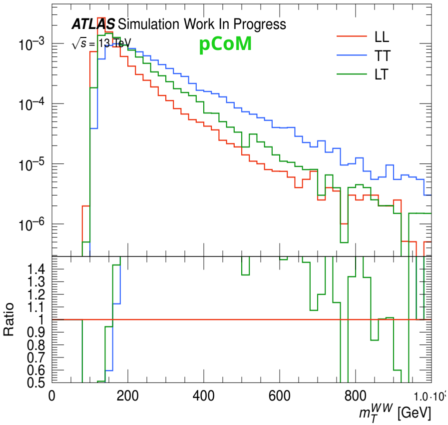
- e channel contributions through γ conversions
- Estimation $\to$ Sherpa 2.2.11 $V\gamma$ MC

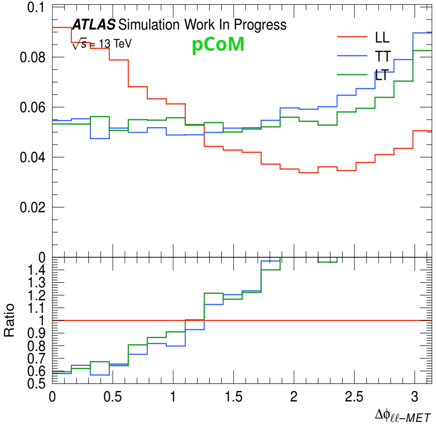
JHEP 2024, 26 (2024)



Various kinematic and angular distributions studied to separate *LL* pol. modes

- Important variables include invariant masses, p_T, $\Delta\phi$, ΔR, etc. of final state particles





(Complete list of distributions in backup slides)

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

Deep Neural Network trained to separate LL-polarization modes

Statistics

Label	Class	No. of events (SR) pCoM	WWCoM
1	Signal	$\sim$ 20k	$\sim$ 22k
0	Background	$\sim$ 250k	$\sim$ 250k

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

Background TL, TT pol., ssWW QCD, ssWW INT, WZ QCD, WZ EW6, ZZ, V/γ, charge Flip, triboson, ttX, Fakes (W+jets, $t\bar{t}$ semi-leptonic)

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

Preprocessing Scaled to MC event weights & standardized normally distributed data

- Distribution with zero mean and unit variance

Scaling Datasets are balanced to avoid bias: scale factor $(Sig) = \sigma(Bkg)/\sigma(Sig)$

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

Avoiding Overtraining Dropout layers, Early stopping, Model checkpoint

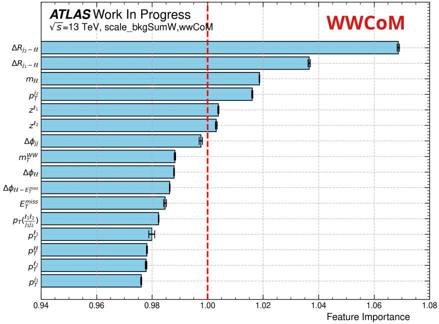
Varible Selection: Permutation Feature Importance

Permutation Feature Importance algorithm:

- Ranks input features of a NN
- Removes highly correlated input variables
- Benefit: Retraining of NN is not needed
- NN evaluated on permuted sets of input features

Method:

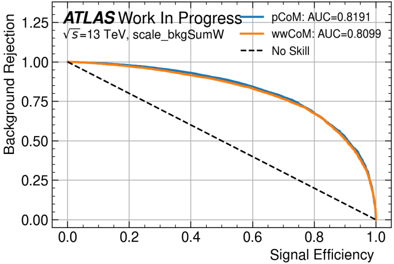
- $\forall$ input features, modified input dataset is created
- Values for this feature (say j) are swapped with each other over the whole dataset
  - It breaks the association of j value to a particular input observable
  - Also breaks j value's correlations to other input observables

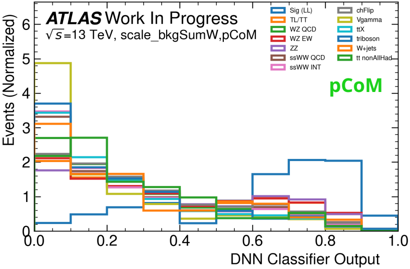


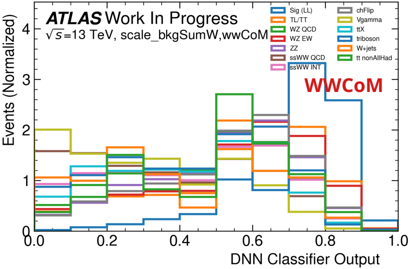
- Algorithm's inputs: trained model f, input dataset X, figure-of-merit L
- Performance measures: $e^{\text{orig}} = L(f(X)),\ e_j^{\text{perm}} = L(f(X_j^{\text{perm}}))$
- Feature Importance: $FI_j = \frac{e_j^{\text{perm}}}{e_j^{\text{orig}}}$

DNN Classification Performance

- Predictions made on Sig & Bkgs using trained DNN model
- ROC curve gives an estimation of the algorithm accuracy
- DNN output score shows a good separation of Sig and individual Bkgs







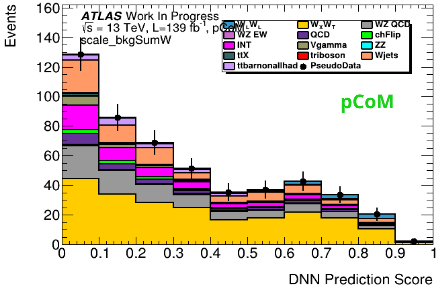
Extracting $W_L W_L$ polarization fraction

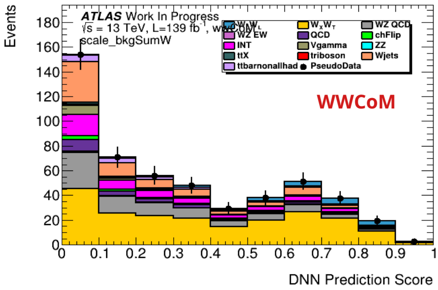
χ^2 fit for $W_L W_L$ polarization fraction (Signal Region):

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

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

$$f_{LL} = \frac{c_{LL} \cdot \sum N_{LL}}{c_{LL} \cdot \sum N_{LL} + \sum N_{XT}}$$





Result*	f_{LL} pCoM	f_{LL} WWCoM
DNN	$0.06^{+0.05}_{-0.06}$	$0.09^{+0.06}_{-0.06}$
Theory prediction	0.0640 ± 0.0007	0.0987 ± 0.0011

Note: Pseudodata$= \sum$ (Sig, all Bkgs)

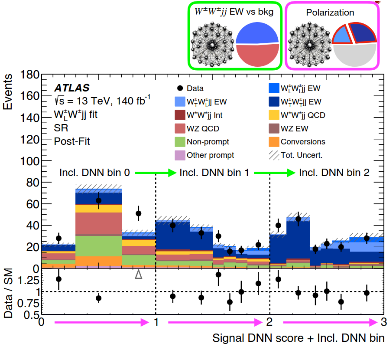
*Systematics not yet included.

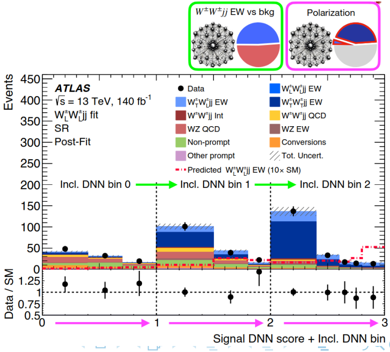
Single boson polarzation ($W_L^{\pm}W^{\pm}jj$):

- Observed (Expected) significance: 3.3(4.0)
- Measured cross-section: 0.88 ± 0.30 fb (in agreement with SM prediction)
- Dominated by statistical uncertainty
- **First evidence** for longitudinal polarization in VBS

Double boson polarization ($W_L^{\pm}W_L^{\pm}jj$):

- Observed (Expected) 95% CL upper limit of 0.45(0.70) fb
- Measured cross-section in agreement with the Standard Model
- Dominated by statistical uncertainty
- **Most stringent limit** for $W_L^{\pm}W_L^{\pm}jj$ to date





Summary

Electroweak polarized *ssWWjj* production:

- Golden channel for VBS
- Polarization studies help in SM validation and BSM searches
- Sensitive to polarization-affecting new physics effects and others like aQGCs

Classification of $W_L W_L$ polarization modes:

- Deep Neural Network constructed to classify $W_L W_L$ vs Bkgs
- $W_L W_L$ polarization fraction (f_{LL}):

Result	f_{LL}	
	pCoM	WWCoM
DNN	$0.06^{+0.05}_{-0.06}$	$0.09^{+0.06}_{-0.06}$
Theory prediction	0.0640 ± 0.0007	0.0987 ± 0.0011

- Current work in progress:
  - Navigating through NAF
  - Update the results with new DNN
  - Including systematics in the fit
  - Training LX vs Bkgs DNN classifier (tentatively)

Thank you!

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
```

$\Longrightarrow$ using explicit τ decays

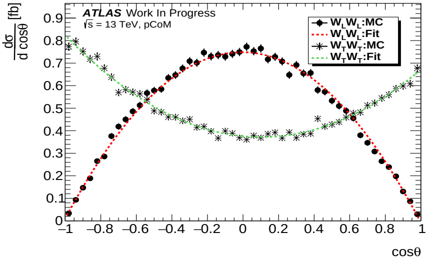
MC Cross-sections: Locally generated samples

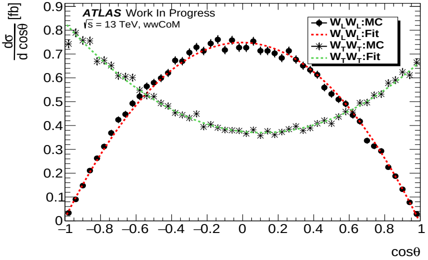
Total events = 100,000 each

Final States $(+jj)$	**BW cutoff**	**Cross-section (fb)** pCoM	WWCoM
$\ell\nu\ell\nu$	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

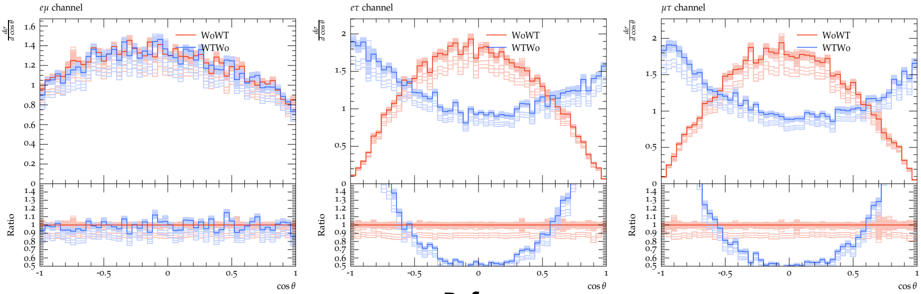
cos θ validation fit:

- Polar angle θ = 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



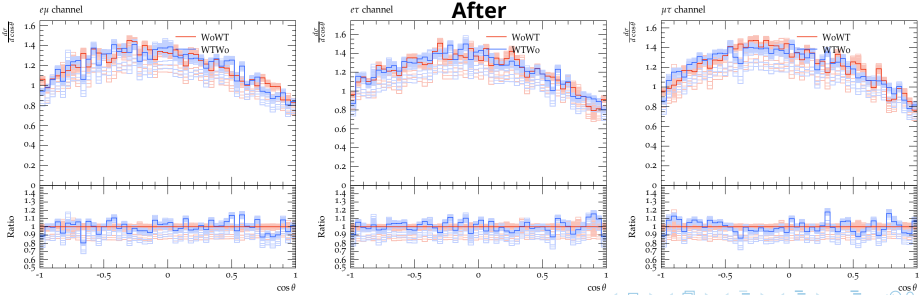


Comparison of new and old $\cos\theta$ plots: Lepton channels



Before

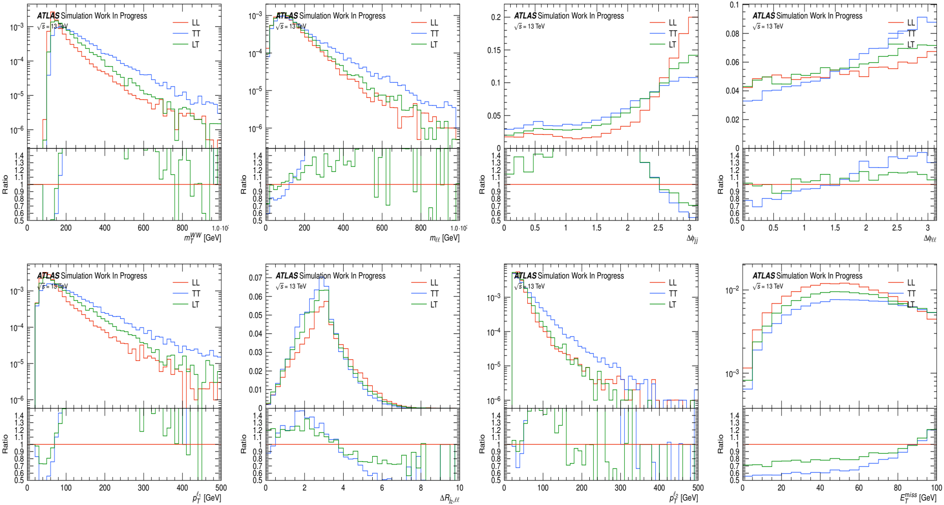
After



Polarization Discriminating Variables

Normalized, pCoM

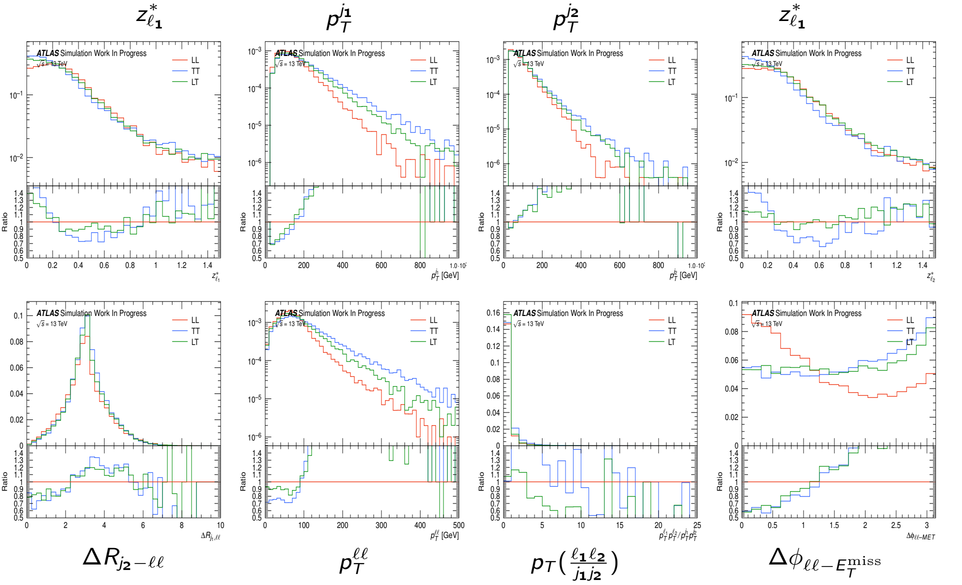
m_T^{WW} | $m_{\ell\ell}$ | $\Delta\phi_{jj}$ | $\Delta\phi_{\ell\ell}$



$p_T^{\ell_1}$ | $\Delta R_{j_2-\ell\ell}$ | $p_T^{\ell_2}$ | E_T^{miss}

Polarization Discriminating Variables

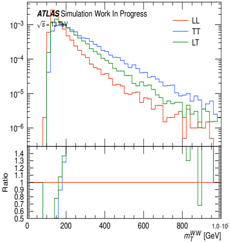
Normalized, pCoM



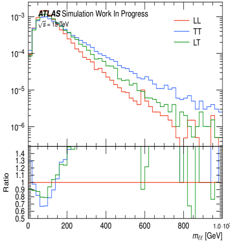
Polarization Discriminating Variables

Normalized, WWCoM

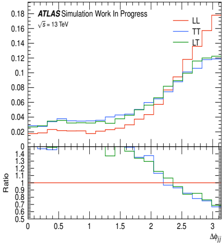
m_T^{WW}



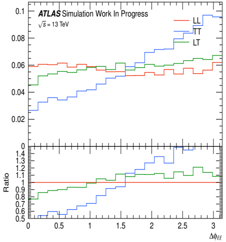
$m_{\ell\ell}$

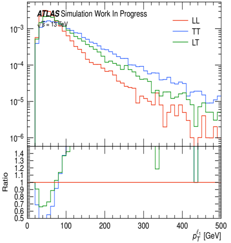


$\Delta\phi_{jj}$

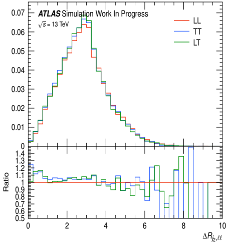


$\Delta\phi_{\ell\ell}$

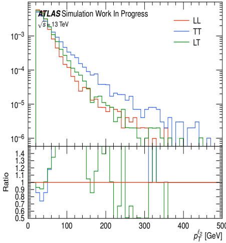




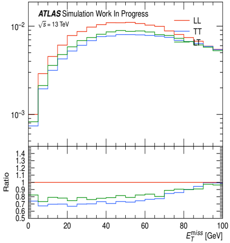
$p_T^{\ell_1}$



$\Delta R_{j_2-\ell\ell}$



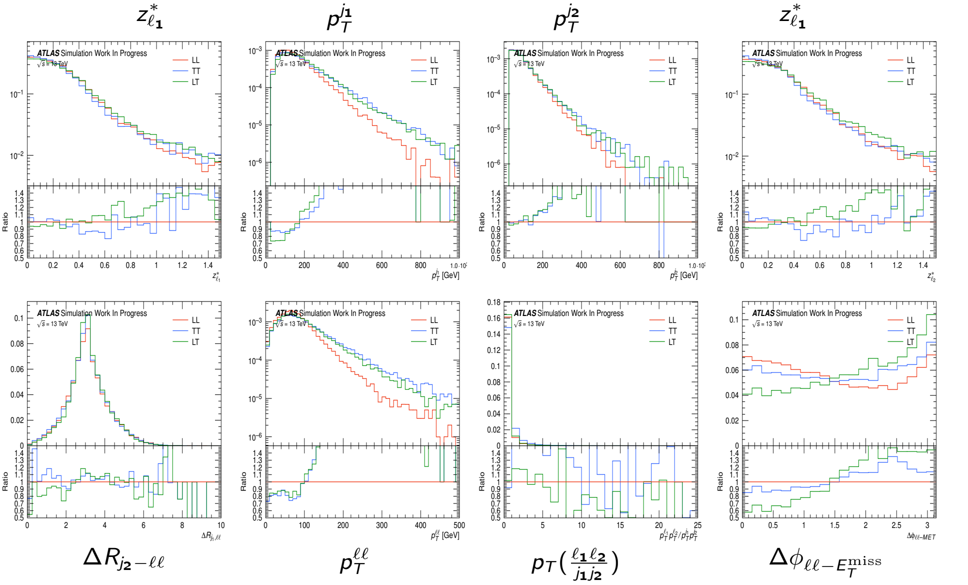
$p_T^{\ell_2}$



E_T^{miss}

Polarization Discriminating Variables

Normalized, WWCoM



Polarization Discriminating Variables

$$m_T^{WW} = \sqrt{(\sum_i E_i)^2 + (\sum_i p_{z,i})^2} \quad (1)$$

$$z^*_{\ell_1} = \left| \frac{\eta_{\ell_1} - 0.5 \cdot (\eta_{j_1} + \eta_{j_2})}{\Delta \eta_{jj}} \right| \quad (2)$$

Hyperparameter Optimization

Hyperparameters: configurations and settings that determine the DNN architecture & training behaviour

Hyperparameter	Optimization Range	Sampling
No. of Hidden layers	$[1, 10]$	In steps of 1
Dropout layers	True/False	Optionally inserted after every hidden layer
Dropout Rate	$[0.0, 0.4]$	In steps of 0.01
No. of Units per layer	$[16, 32, 64, 128, 256, 512]$	In exponents of 2
Learning Rate (Adam)	$[10^{-2}, 10^{-5}]$	Logarithmic
Batch Size	$[16, 32, 64, 128, 256]$	In exponents of 2

- Values are optimized such that the validation loss is minimized
- Improves the precision and accuracy of model
- Some values are set by manual tuning, and the rest using `Optuna`
- Chronology for DNN Optimization:
  Hyperparameter Optimization 1 $\rightarrow$ Input Feature Reduction $\rightarrow$ Hyperparameter Optimization 2

Variable Selection: Permutation Feature Importance

Permutation Feature Importance algorithm:

- Ranks input features of a NN
- Removes highly correlated input variables
- Benefit: Retraining of NN is not needed
- NN evaluated on permuted sets of input features

Method:

- $\forall$ input features, modified input dataset is created
- Values for this feature (say j) are swapped with each other over the whole dataset
  - No new values created
  - It breaks the association of j value to a particular input observable
  - Also breaks j value's correlations to other input observables

Variable Selection: Permutation Feature Importance

- Algorithm's inputs:
  1. trained model f
  2. input dataset X
  3. figure-of-merit L
- Original performance measure: $e^{\mathrm{orig}} = L(f(X))$
- Permuted performance measure: $e_j^{\mathrm{perm}} = L(f(X_j^{\mathrm{perm}}))$
- Model Reliance/Feature Importance: $FI_j = \frac{e_j^{\mathrm{perm}}}{e_j^{\mathrm{orig}}}$ or $e_j^{\mathrm{perm}} - e_j^{\mathrm{orig}}$
  - $FI = 1 \implies$ no reliance on X_1
  - $FI = 2 \implies$ loss doubles when X_1 is scrambled, heavy reliance on X_1
  - $FI < 1 \implies$ rely less on X_1 than a random guess, difficult to interpret

Error estimation on FI:

- Shuffling of X performed 10 times
  - Mean, std dev $\rightarrow FI_j, \Delta FI_j$

DNN Model Details

Model:

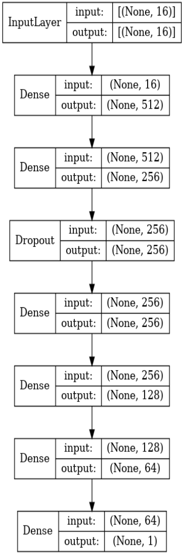
- 5 hidden layers: random normal init., ReLU activation
- 2 dropout layers (40%) to reduce overtraining
- 1 sigmoid output layer with random uniform activation

Compilation:

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

Model Fit:

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

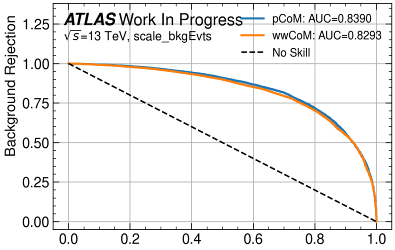


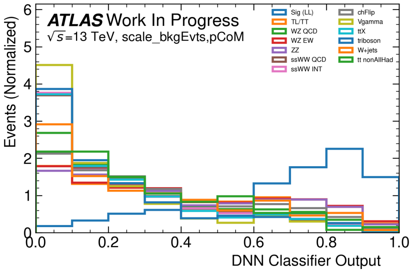
nEvts and Sum of Weights

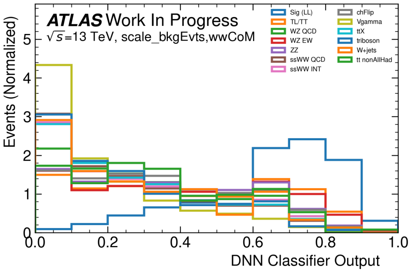
Process	No. of Evts	Sum of Weights
LL (pCoM)	19673	12.9161
LL (WWCoM)	21643	20.6243
XT (pCoM)	67791	140.759 (TT) + 76.8323 (LT)
XT (WWCoM)	68658	142.854 (TT) + 67.2892 (LT)
Diboson: WZ QCD	27800	82.753
Diboson: WZ EW	2515	4.0837
Diboson: ZZ	2880	2.50948
ssWW QCD	25006	22.9733
ssWW INT	114209	48.3424
Fakes: W+jets	151	71.0203
Fakes: $t\bar{t}$ non all hadronic	133	16.933
chFlip	8231	10.0955
V/γ	614	12.5289
ttX	2168	4.03342
triboson	308	0.649666
Bkg (pCoM)	251806	493.513966
Bkg (WWCoM)	252673	486.065866

Scale:

$$class\_W(sig) = \frac{n(Bkg)}{n(Sig)}$$

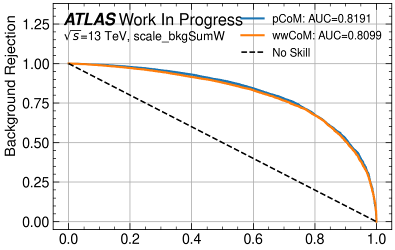


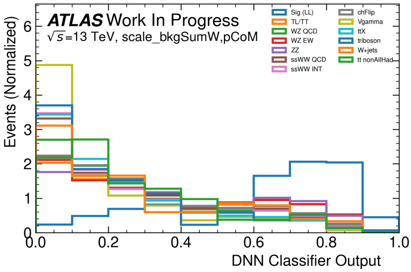


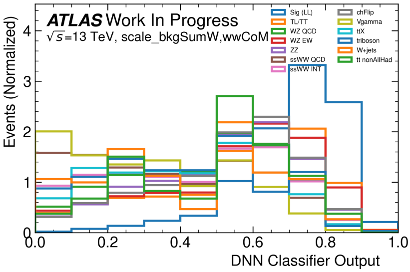


Scale:

$$event\_W \times \frac{sumW(Bkg)}{sumW(Sig)}$$







DNN Discriminant for Classification

pCoM, WWCoM