VBS opposite-sign WW in ATLAS

Multi-Boson Interactions 2023

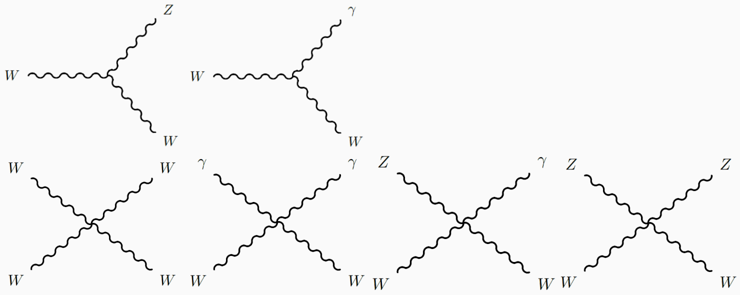
Christophe Roland

August 31, 2023

LPNHE/CNRS
On Behalf of the ATLAS Collaboration

The Standard Model and the electroweak interaction

- The electroweak theory provides a unified description of the electromagnetic and weak interactions.
- After symmetry breaking, there are three massive vector bosons $W^{\pm}, Z$ (weak interaction) and one massless vector boson γ (the photon). They are interactions terms in the Lagrangian corresponding to interactions between these particles:



$\rightarrow$ Measurement of vector boson fusion and vector boson scattering processes in the LHC can be used to probe these interactions.

Vector boson scattering

- Let's consider the scattering of two opposite-sign longitudinally polarized W bosons:

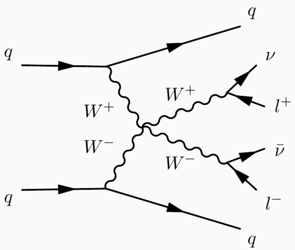
$$W_L^+ W_L^- \longrightarrow W_L^+ W_L^-$$

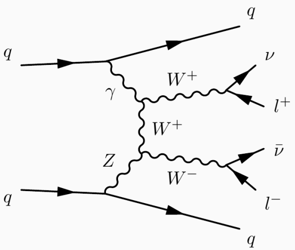
  Computing the cross-section without taking the Higgs into account leads to a violation of unitarity at the $\sim$ TeV scale. In the Standard Model, the Higgs boson is sufficient to completely restore unitarity.
- But some BSM (beyond the Standard Model) models predict that the Higgs only partially unitarizes this cross-section.
- For example, in two-Higgs doublet models, there are two CP-even Higgs bosons, the lighter one being identified with the 125 GeV Higgs boson discovered in 2012. The lighter boson would only partially unitarize the cross-section while the heavier one would complete the unitarization at higher energy. The WW scattering cross-section would then be larger than predicted by the Standard Model.

$\rightarrow$ Important to test the Standard Model predictions for VBS processes!

Overview

- First observation of electroweak $W^+W^- +$ jets production in ATLAS !
- Probe the EWK symmetry breaking of the SM, process sensitive to quartic gauge coupling.
- Two neural networks are trained to distinguish our signal from our largest backgrounds: top and QCD WW production.
- We are looking at events where the two W bosons decay leptonically, the final state is either two or three jets, two leptons (one electron and one muon) and missing transverse energy.





Strategy

- We define one signal (further split into a 'two jets' and a 'three jets' region) and one control region to constrain the top uncertainties.
- The two regions are identical except that we apply a b-veto in the signal region but require that one of the two leading jets be a b-jet in the control region.
- The cuts are fairly loose but we train and use two neural networks (one for $n_{jets} = 2$ and another for $n_{jets} = 3$) to have an optimal signal/background separation.
- A validation region is defined for low (< 0.6) neural output (i.e. small signal and large background) to check the correlations between the variables used by the neural network among themselves and with the neural network output by comparing with data.
- The neural network outputs are used to perform a likelihood fit using the maximum likelihood estimation method. From the fit, the signal strength (the ratio between the measurement and the expected cross-section) is inferred.

Selection

Selection (signal and control region):

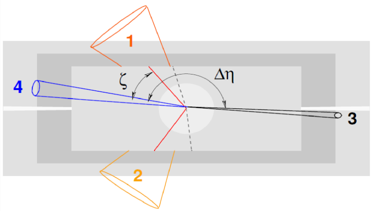
- Two opposite sign tight isolated leptons with $p_T > 27$ GeV (one electron and one muon)
- third lepton veto ($p_T > 10$ GeV, tight if $p_T < 25$ GeV and medium otherwise)
- $p_T^{miss} > 15$ GeV
- two or three jets with $p_T > 25$ GeV, $|\eta| < 4.5$
- centrality $= \min\left[\min(\eta^{lep0}, \eta^{lep1}) - \min(\eta^{jet0}, \eta^{jet1}),\right.$ $\left.\max(\eta^{jet0}, \eta^{jet1}) - \max(\eta^{lep0}, \eta^{lep1})\right] > 0.5$ (was found to improve neural network performance if applied before training)
- $m_{ll} > 80$ GeV (to suppress VBF HWW)
- b-jets (b-tagging with the 85% working point):
  - b-jet veto (signal region)
  - one of the two leading jets is a b-jet (control region)

Centrality

- Centrality is defined by:

$$\zeta = \text{centrality} = \min\{[\min(\eta_{\ell_1}, \eta_{\ell_2}) - \min(\eta_{j_1}, \eta_{j_2})],\\ [\max(\eta_{j_1}, \eta_{j_2}) - \max(\eta_{\ell_1}, \eta_{\ell_2})]\}$$

Vector boson scattering events tend to have a larger value of ζ.



- For three jets, third jet centrality is also used:

$$\text{Third jet centrality} = \left| y_{j_3} - \frac{1}{2} \times \frac{y_{j_1} + y_{j_2}}{y_{j_1} - y_{j_2}} \right|$$

The centrality of the third jet in signal events peaks at high values as it is frequently emitted from one of the two leading forward jets.

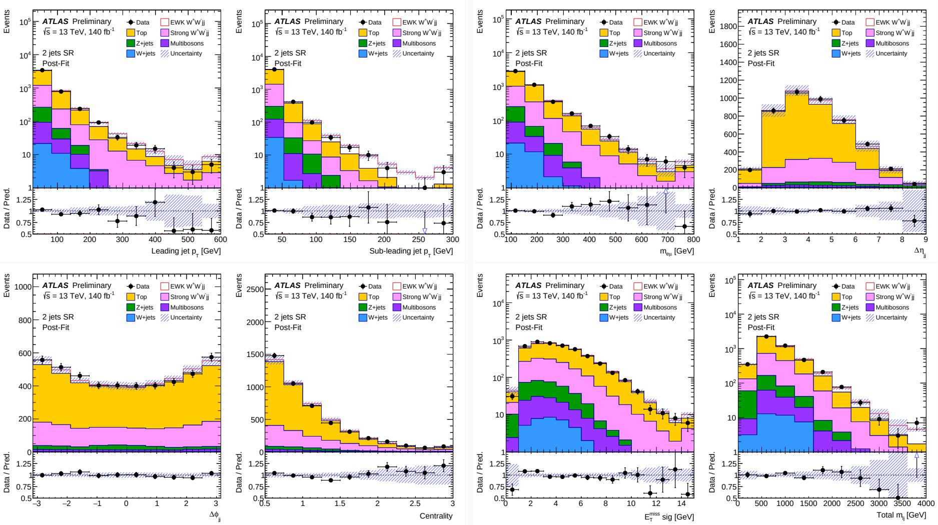
Neural networks

- We train two different networks: one for two jets and one for three jets, using TMVA and Keras.
- Variables used for training are: leading and sub-leading jet p_T, m_{ll}, $\Delta\eta_{jj}$, $\Delta\varphi_{jj}$, centrality, E_T^{miss}significance and total m_{lj}.
- For three jets: third jet p_T and third jet centrality are also included.
- To avoid overtraining of the NN and increase its robustness, a dropout regularisation with a rate of 0.1 is applied.
- The NN was validated in the signal region by comparing data and simulation for the correlations of the input variables between each other and against the NN output.
- Two hidden layers are used with 108 nodes on the first hidden layer and 60 on the second one.

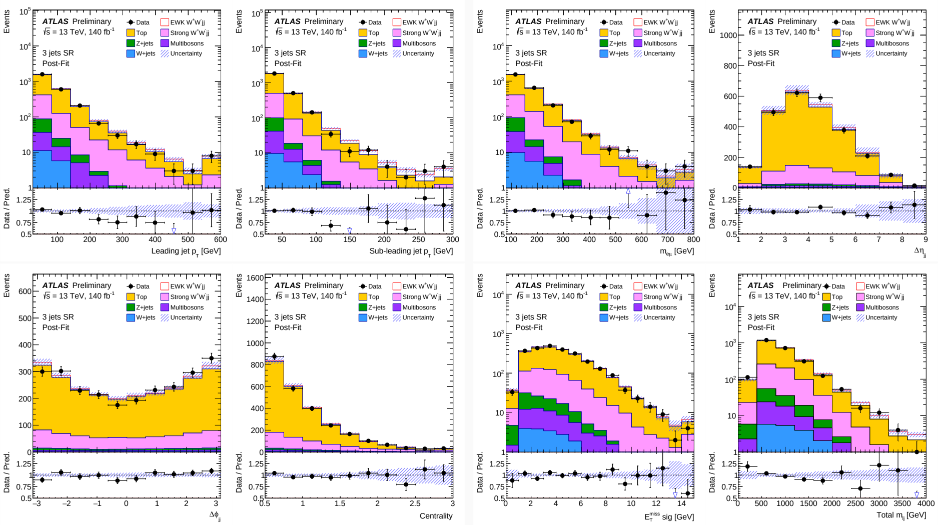
Observables in the signal region

- Next two slides: plots in the signal region without cut on the neural network output, for two and three jet events.
- We can see a good agreement between SM prediction and data, so the input training variables are well modeled in the signal and control region.
- The plotted observables are the variables used by the neural network.

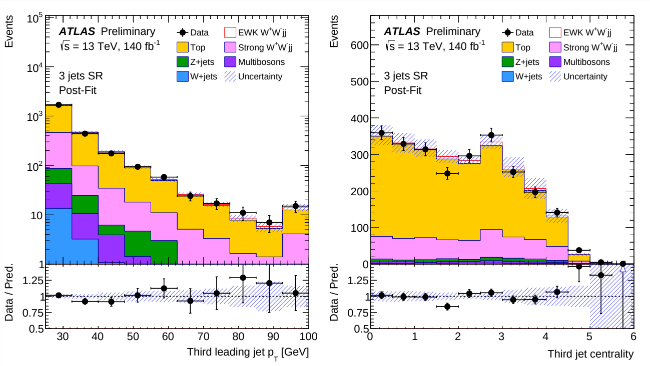
Signal region plots - NN training variables - two jets



Signal region plots - NN training variables - three jets



- Variables for the third jet:



Uncertainties

The uncertainties considered are the following:

- Experimental uncertainties: the largest are the jet energy scale uncertainties, the b-tagging efficiency, the jet flavor composition uncertainty and the jet energy scale dependence on pile-up.
- Theoretical uncertainties:
  - Signal: PDF, scale, interference, initial and final state radiation uncertainties.
  - Top: PDF, Matrix element, parton shower, scale, initial and final state radiation uncertainties, $Wt/t\bar{t}$ interference.
  - QCD-WW and Z+jets: PDF and scale uncertainties.
- Monte Carlo and data statistical uncertainty.

Fake lepton background

- Due to its very small contribution, the background consisting of W boson production in association with jets is modeled by simulation, yet further constrained from data in a region enriched in fake leptons to better account for the fraction of events where jets are reconstructed as a lepton.
- Leptons with loose identification and by requiring that at least one of the two leptons fails the isolation and the identification requirements but otherwise the same as the signal region.

Likelihood fit

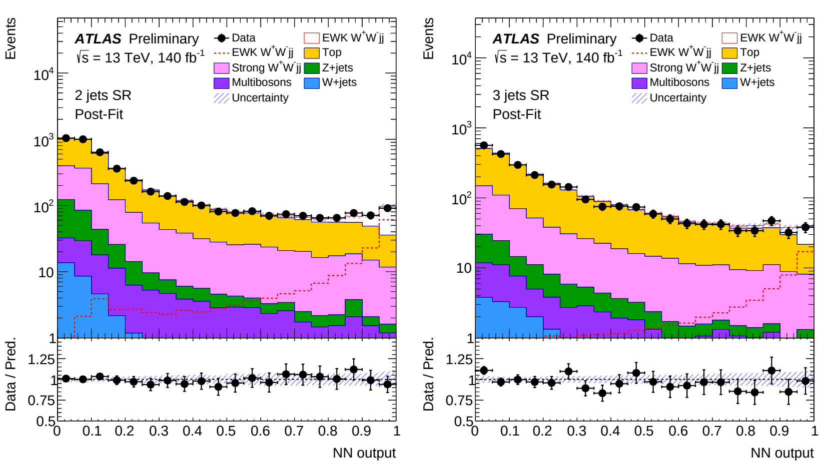
A likelihood fit of the neural network output distributions is performed using the TRExFitter framework. The likelihood function is the probability density to observe $n_1, n_2, ...$ in the first, second, ... bin of the NN output distribution:

$$L(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{n_{\text{bins}}} p(n_i|\mu, \boldsymbol{\theta})$$

$-\ln(L(\mu, \boldsymbol{\theta}))$ is then minimized to obtain the best-fit value for the signal strength $\hat{\mu}$.

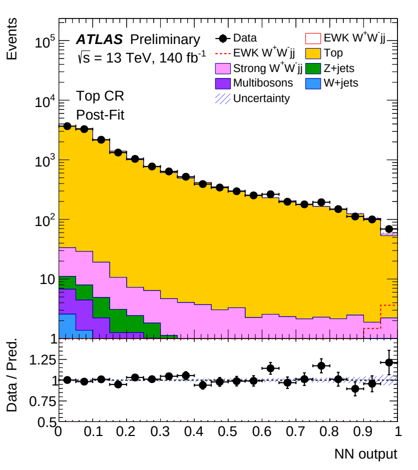
- There is a pruning used to remove negligible uncertainties with a threshold of 0.5%.
- The top and QCD-*WW* normalization are floating parameters in the fit.

Neural network output distribution in the signal region



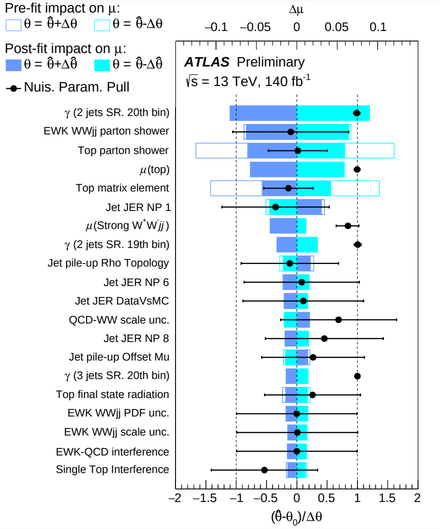
- Left: 2 jets, right: 3 jets.
- Neural network distribution is fairly well modeled.
- The observed (expected) significance is 7.1σ (6.2σ), for both 2 and 3 jets combined.

Neural network output distribution in the control region



- Neural network distribution is fairly well modeled in the control region, but with some fluctuations.

Nuisance parameters ranking



- The largest systematics are shown here, ranked by post-fit impact.
- γ corresponds to the MC statistical uncertainty.
- The matrix element unc. and Parton shower unc. are for the top background (by comparing Powheg with MadGraph and for the showering: Pythia with Herwig).
- The signal parton shower uncertainty is estimated by comparing Powheg+Pythia (nominal) with Powheg+Herwig.

Summary of the uncertainties

Systematic	$\frac{\sqrt{(\Delta\mu)^2-(\Delta\mu')^2}}{\mu}$
Monte Carlo statistical uncertainty	7.67%
Top quark theoretical uncertainties	6.30%
Signal theoretical uncertainties	5.80%
Jet experimental uncertainties	4.92%
Strong WWjj theoretical uncertainties	1.33%
Luminosity	0.83%
Fake uncertainty	0.48%
b-tagging	0.39%
Lepton experimental uncertainties	0.14%
Others	0.31%
Statistical uncertainty	12.3%
Top normalization uncertainty	4.90%
QCD-WW normalization uncertainty	2.24%
Total uncertainty	18.5%

- We are dominated by the data statistical uncertainty. It is estimated by running the fit again while ignoring all the other sources of uncertainties.

Defining the fiducial region

The cross-section is measured in a fiducial region that is designed to be similar to the part of the signal region which is most sensitive to the signal cross-selection $\hat{\mu}$. It is estimated by multiplying the predicted cross-section by the signal strength. A cut on m_{jj} is added in the fiducial region definition:

1. It is ambiguous whether tribosons processes corresponding to Feynman diagrams with electroweak order equal to six ought to be defined as part of our signal or not. We are using POWHEGBOX and this tribosons contribution is not included, but the same process simulated using MadGraph would include it by default. Adding a cut on m_{jj} strongly suppresses this contribution and makes comparison with other generators easier.
2. Our signal region being fairly inclusive, the signal strength obtained from the fit is not very sensitive to signal events with low neural network output values. So, the point of cutting on m_{jj} in the fiducial region definition is to make sure that we are avoiding a potential over-extrapolation.

Fiducial region

Requirements
One electron and one muon with opposite charges
No additional lepton
$p_T^{\text{dressed } \ell} = p_T^{\ell} + \sum_i p_T^{\gamma_i}$ if $\Delta R(\ell, \gamma_i) < 0.1$
Leptons p_T: $p_T > 27$ GeV
Lepton η: $\vert\eta\vert < 2.5$
Two or three jets with $p_T > 25$ GeV and $\vert\eta\vert < 4.5$
$p_T^{miss} > 15$ GeV
$m_{jj} > 500$ GeV
centrality > 0.5
$m_{e\mu} > 80$ GeV
b-jet veto

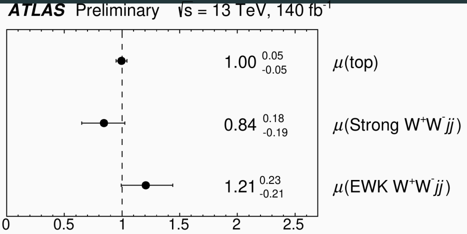
- The fiducial region is defined with a selection that is similar to the reconstructed signal region, with an additional cut on m_{jj}.

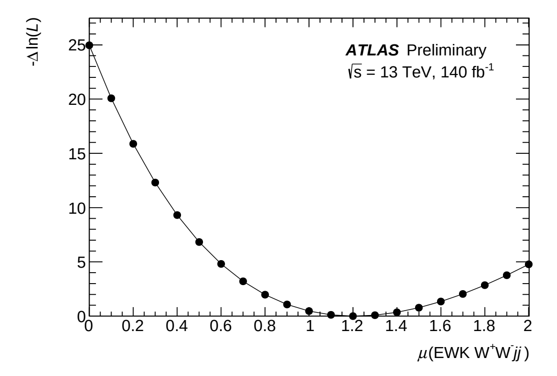
Cross-section

The signal strength and cross-section are:

$\mu = 1.21^{+0.23}_{-0.21}$

$\sigma_{\text{observed}} = 2.65^{+0.52}_{-0.48}$ fb





Conclusion

- This is an observation of the electroweak production of two opposite-sign W bosons with two or three jets, using data collected with the ATLAS detector, during the LHC run-II corresponding to a total integrated luminosity of 140 fb^{-1}. This is the first measurement and first observation of this process in ATLAS.
- The result is in agreement with Standard Model predictions, with an observation significance of 7.1σ.
- $\rightarrow$ ATLAS-CONF-2023-039

BACKUP

Event Yield

	Event yields	
Process	$n_{\text{jets}} = 2$	$n_{\text{jets}} = 3$
EWK W^+W^-jj	158 ± 27	54 ± 13
Top quark	2885 ± 214	1851 ± 131
Strong W^+W^-jj	1214 ± 256	514 ± 121
W+jets	37 ± 97	19 ± 48
Z+jets	216 ± 62	65 ± 25
Multiboson	101 ± 5	42 ± 3
SM prediction	4610 ± 77	2546 ± 48
Data	4610	2533

Table 1: Number of data events in the signal region compared to the yields from the prediction of the Standard Model. The composition of the latter is also provided, and the events are split in two categories, depending on the number of jets. The uncertainties include both statistical and systematic contributions, and correspond to the values after the likelihood fit.

m_{jj} cut

In order to determine the most appropriate value for fiducial region cut on m_{jj}, we look at three criteria:

- The generator-level cut on m_{jj} must reduce the number of reconstructed events in low NN bins, i.e. where signal/background is very small ($< 5\%$ which corresponds to NN < 0.6) but not significantly reduce the number of events in high NN bins (NN > 0.6).
- The m_{jj} generator-level cut efficiencies must be similar to the NN > 0.6 reconstruction cut efficiencies because we want the fiducial region to be similar to the NN > 0.6 reconstructed region.
- The m_{jj} cut must suppress tribosons contributions.

Choosing the m_{jj} cut at 500 GeV satisfy all three criteria.