The Search for Evidence of Vector Boson Scattering Between a W boson and a Photon in Proton-Proton Collisions at the L.H.C.

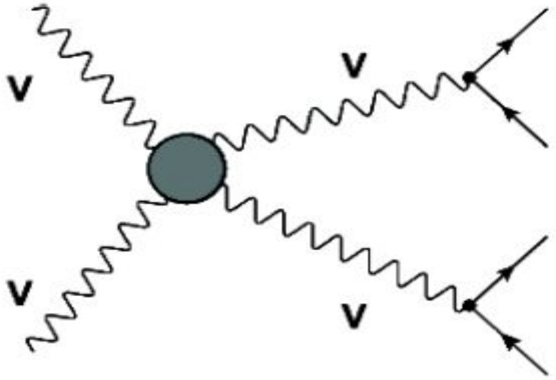
John McGowan, McGill University
Supervisor: Prof. Brigitte Vachon

McGill

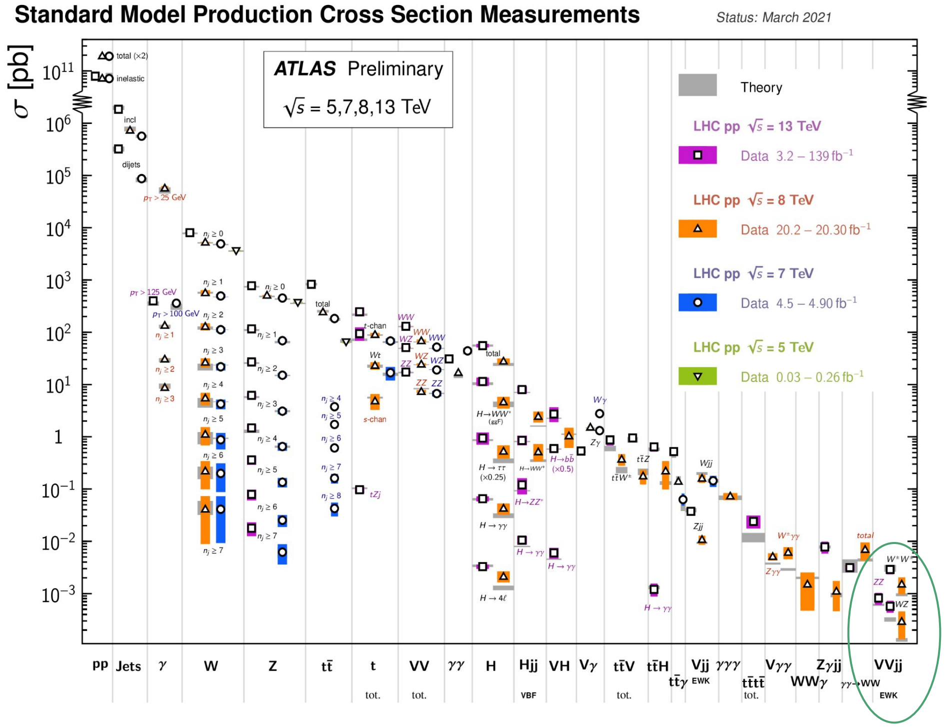
ATLAS
EXPERIMENT

Why Study Vector Boson Scattering (VBS) at the LHC?

- In the Standard Model, the interactions between gauge bosons are completely specified by the SU(2) x U(1) structure of the theory.
- This makes the study of the interactions between gauge bosons a powerful approach to search for new physics.
- Any deviation from SM predictions in the self interactions of gauge bosons would indicate the presence of new physics phenomena.



VBS at the LHC



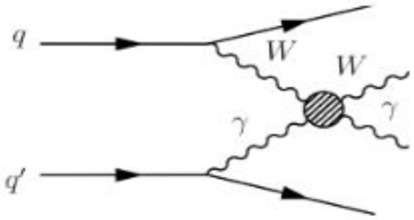
- The LHC provides a unique environment in which to study rare Standard Model processes.
- The search for evidence of scattering between a W boson and a photon is carried out using a total of 139 fb^{-1} data collected by the ATLAS detector at $\sqrt{s}$ = 13 TeV.

ATL-PHYS-PUB-2021-005

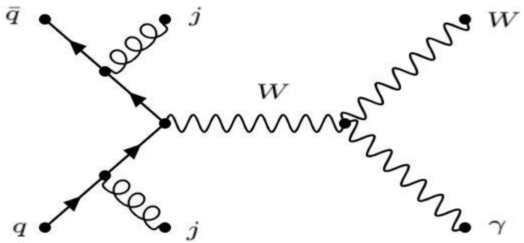
Challenges of VBS Measurements

- The search for the scattering of a W boson and a photon comes with formidable challenges.
  - There is a large and irreducible background from poorly modelled processes involving strong interactions between particles (QCD):

Signal



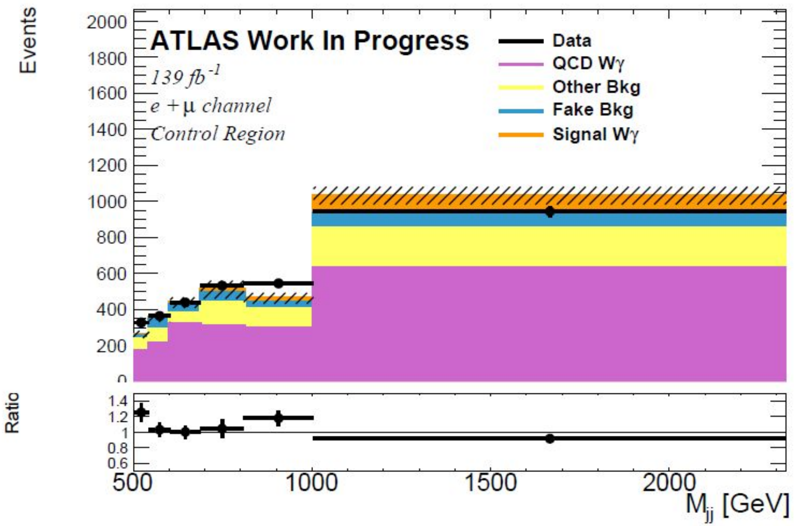
Background

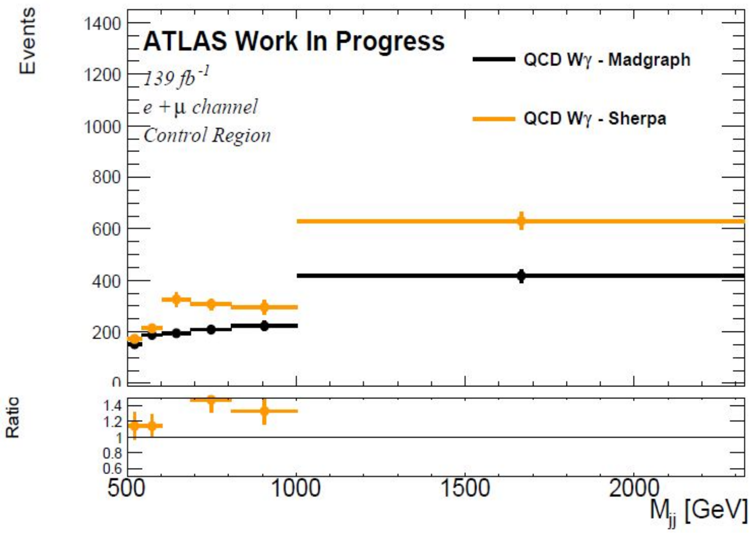


  - The imperfect modelling of the detector response results in a non-negligible number of jets being misidentified as photons.
- In this talk I will discuss two approaches to estimating the number of signal events from the large irreducible background.

Method 1: Fitting Both QCD and Signal Contributions to Data

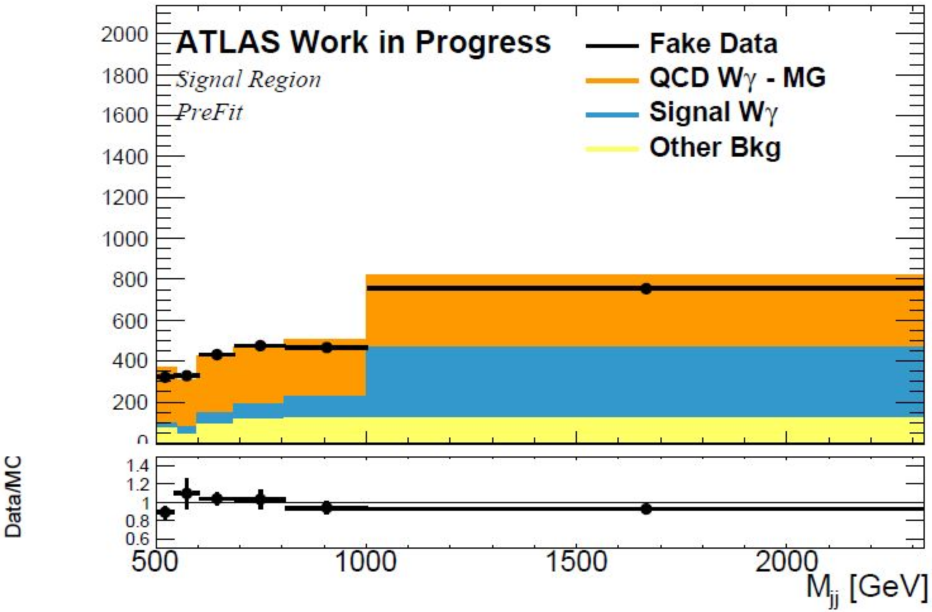
- To constrain the QCD background in a signal region, we can fit the difference between data and MC in a QCD enhanced control region, and apply a reweighting to QCD simulations in the signal region.
- Simultaneously fit the electroweak component.

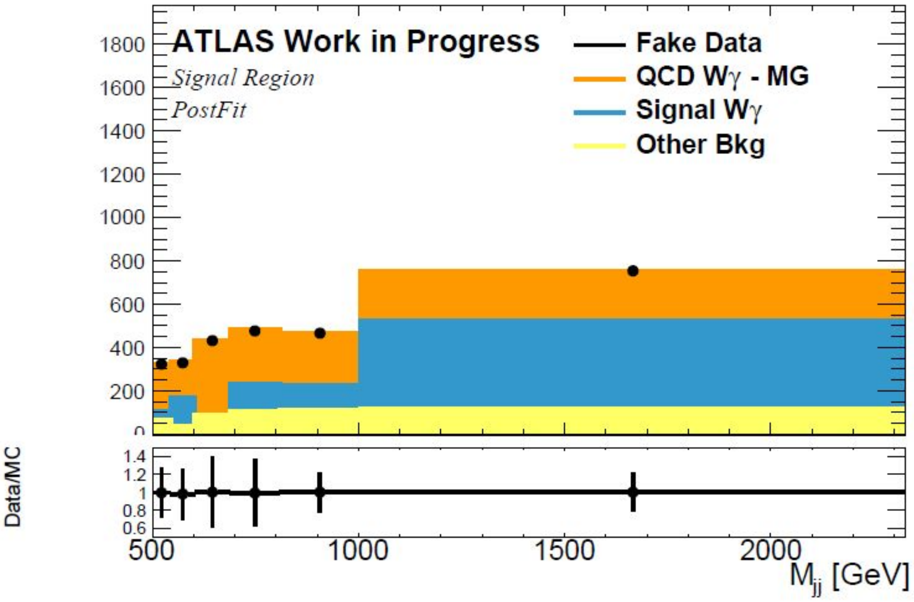




Method 1: Checking The Method With Different Simulations

- To check this method without unblinding in our signal region, we can use a fake dataset made up of simulated data, with the QCD simulation coming from a different generator.

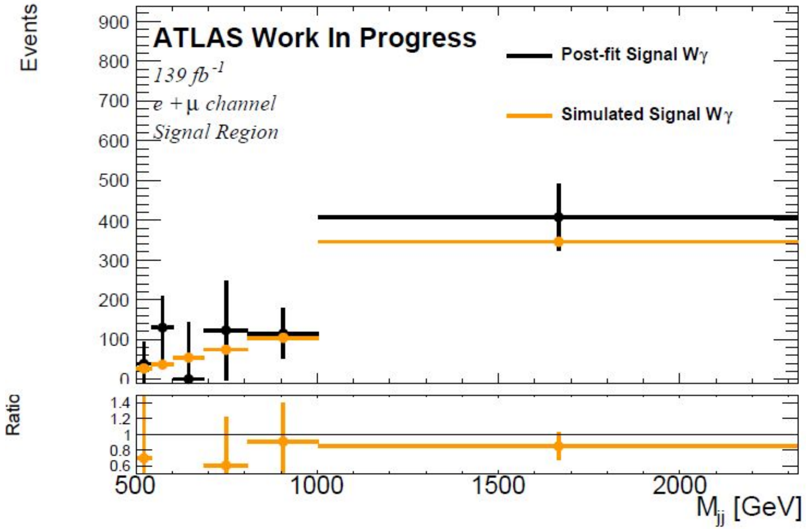


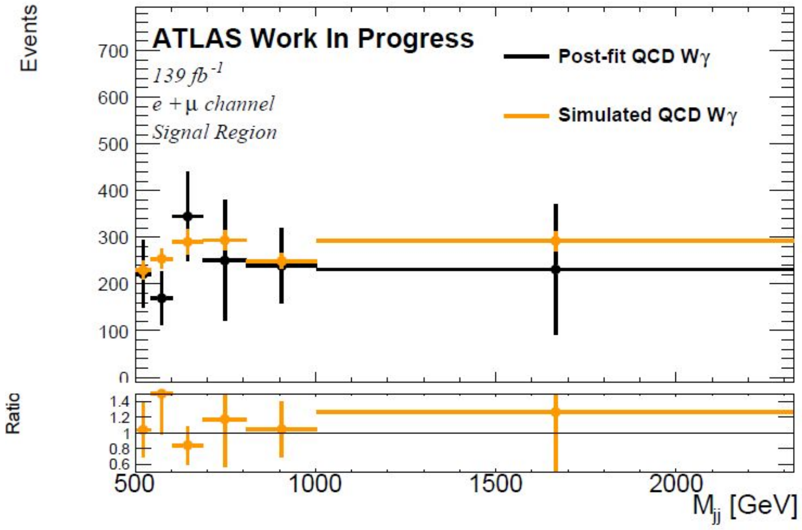


- Shown is the simulated templates compared with the simulated data with a different QCD simulation before and after fitting.

Method 1: Post Fit Distributions

- Now can look and see if we correctly fit the signal and QCD components of the fake dataset.

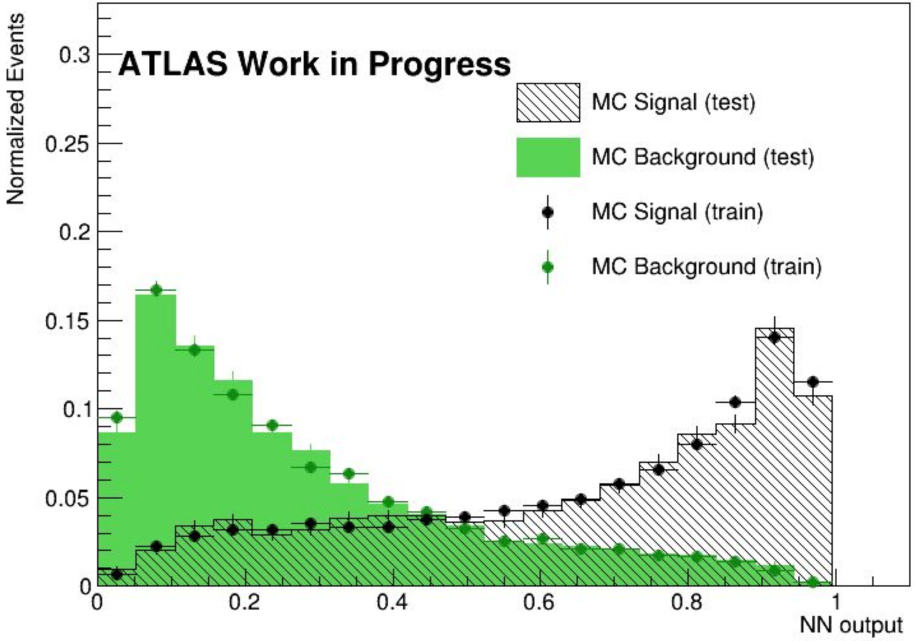




- The agreement is quite good. But limited by statistical error in the control regions.

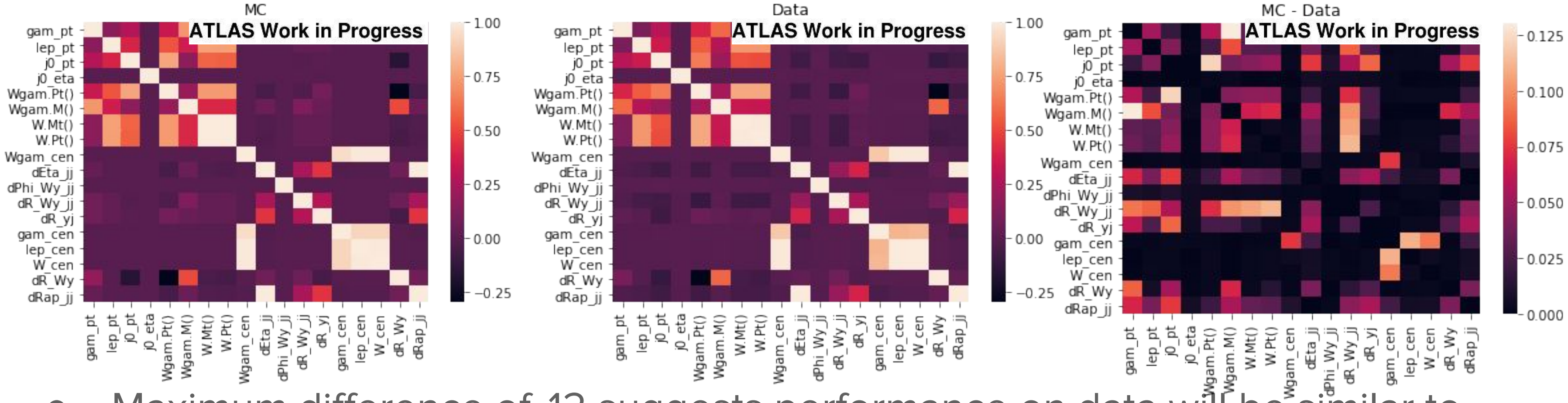
Method 2: Distinguishing Electroweak Signal from Irreducible Background Using a Neural Network

- Another approach that is not limited by statistics in a control region is to train a neural network.
  - 5 layer fully connected neural network.
  - Trained on 18 variables.
  - 50% of simulated sample used as training set, 50% used as testing and validation.
  - 2 NNs trained so all simulated data is used. Average of two networks will be used on data.
- Leverages the discriminating power of multiple variables
- NN distinguishes signal from QCD background with 69% accuracy.
  - Signal defined as event with NN score > 0.5.



Method 2: Generalizing Performance from MC to Data

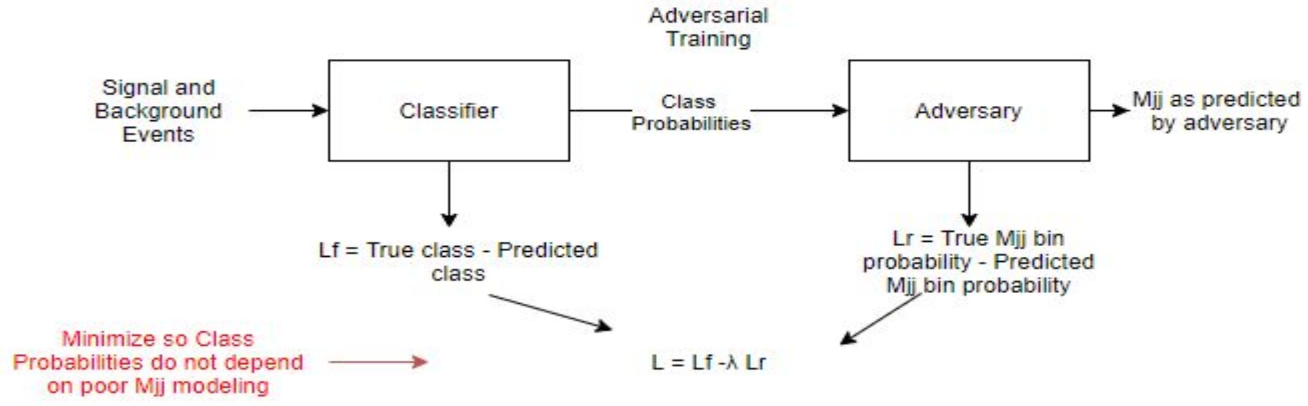
- The neural network will learn correlations between training variables. Similar performance can be expected when evaluating the model on data if the correlation matrices are similar.



- Maximum difference of .12 suggests performance on data will be similar to performance on simulated data.

Next Steps: Adversarial Training

- Irreducible QCD background is mismodelled at high values of di-jet invariant mass (M_{jj})
- Train a second neural network to learn the Mjj distribution (predict M_{jj} bins) from the output of the classifier.



Summary

- VBS measurements provide a powerful probe of new physics.
- The search for the vector boson scattering of a photon and a W boson comes with the significant challenge of a large irreducible QCD background.
- It is possible to constrain the QCD background using a fit which fits both the QCD and electroweak component of the data.
- A neural network has been trained to strongly discriminate the electroweak signal from the dominant QCD background.
- New techniques being explored for regularizing the neural network to not learn MC mismodelling.