SUPER program final report
VLQ Wq+X

Yarelis D. Acevedo Rios

1 Introduction

The analysis on VLQ Wq + X is searching for pair production of a new heavy quark (Q) that decays into a W boson and a light quark (q) in the final state where one W boson decays leptonically (to an electron or muon plus a neutrino) and the other W boson decays hadronically. [Col] Using data collected by ATLAS detector we are looking for new Heavy Quarks (Q) that couple to light SM quarks and W-bosons. In order to analyze the collected data the team generated ROOT files from Ntuples containing kinematic variables of particles. The variables of interest are Jet pT, Lepton pT, MET, $\Delta\phi$(Q, Q),ΔR(lep, ν), and ST. On the analysis, we have a selection criteria to selected events. Cutting for events to maximize the signal efficiency. Reducing the cut will result in more signal, but also more background. The main goal is the determination of what cut values use to maximize the sensitivity, and keep as much signal as possible.

2 Background

The vector like quarks (VLQ) are hypothetical particles that appear in some extensions on the Standard Model (SM). These hypothetical particles are proposed in order to correct shortcomings of SM. VLQs are predicted to decay to a heavy SM boson (W, Z or Higgs) and a quark. The ratio of each decay depends on the model.

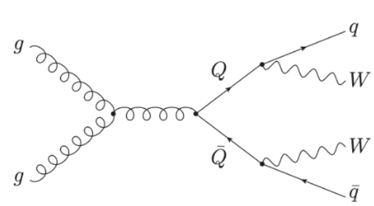

Figure 1: VLQ signal

If such quarks exists they are expected to be produced in pairs. This signal can be similar to many SM processes. The most dominate SM process is the

W+jets.

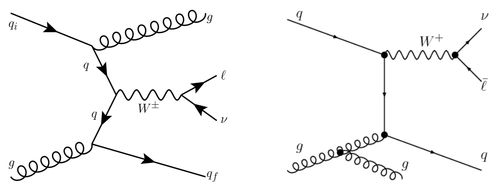

Figure 2: W+jets

3 Setup

At the beginning of this project, I learned how to setup up my personal computer to connect to lxplus for the first time and vncserver. Using Joe Haley's tutorial to setup the vncserver helped. Other things that I learned in the way to be able to connect to the framework was Git and their commands, how to use texts editors more efficiently such as EMACS and NANO. Before getting my hands on the framework, I did a branch for me in order to work on my area.
The framework is written in C++ and Python and uses ROOT to analyze the N-tuples. I had some background on C++ and on Python but not in ROOT. Slowly I learned how the framework worked and how the histograms were generated. Most importantly, how the memory was being handled in the process.

4 Analysis

We are looking for events with two vector like quarks (VLQ):

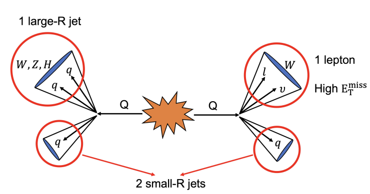

Figure 3: VLQ decay

One VLQ decays to a hadronically decaying W-boson and a light quark. The other VLQ decays to a leptonically decaying W-boson and a light quark. This produces a lepton, missing momentum, a large R-jet from the hadronically decaying W boson, and two small R-jets in the detector.

The main goal is to keep as much signal while removing SM processes. In order to do it we apply selection criteria (cuts) to each event. Evan Van de Wall, the doctoral student of Joe Haley, decided the baseline selection requirements that needed to be optimized. I optimized the selection for the Leading Jet pT and Lepton pT. The optimization was done for VLQs with mass of 1400 GeV. The SM processes that have similar kinematics came from W+jets, ttbar, Z+jets, Single top, Diboson and Multijet events.

The potential cuts values used for the Leading Jet pT are starting from 200 GeV – 600 GeV. For the Lepton pT are Starting from 60 GeV – 120 GeV.

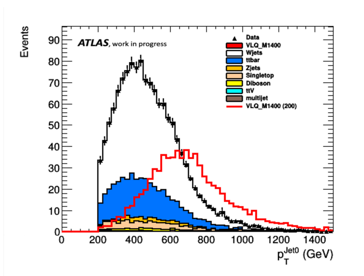

Figure 4: Leading jet pT

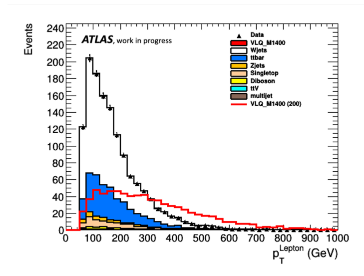

Figure 5: Lepton pT

Figure 6: Potential cut values

After each potential cuts, checked the sensitivity vs number the of signal events. As mentioned before, the higher number of sensitivity and signal events the better.

5 Results

The signal sensitivity was calculated by taking the signal (represented as S) over the square root of the signal and the background (represented as B).

The significance plots for Leading Jet pT 200 – 600 GeV/Lep pT < 60 GeV:

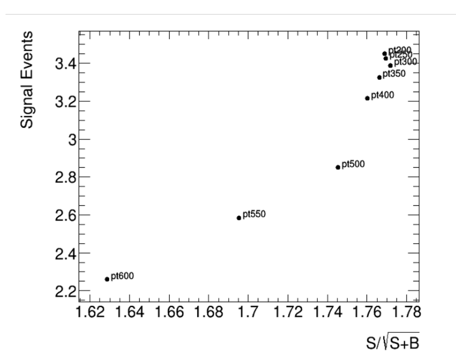

Figure 7: Significance plot for Leading jet pT

Each point is the cuts made for the Leading jet pT from 200 GeV – 600 GeV in steps of 50. As the Leading jet increases the energy, it loses events so the best cut is between 200 – 300 GeV, as observed on Figure 7.

Significance plots for Leading Jet pT < 200 GeV / Lep pT < 60 - 120 GeV

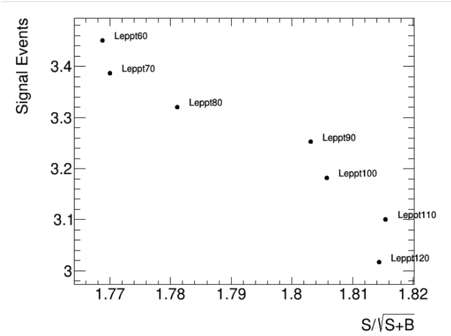

Figure 8: Significance plot for Lepton pT

Each point is the cuts made for the Lepton pT from 60 GeV – 120 GeV in steps of 10. As the Lepton pT increases, it has more sensitivity, but it loses a lot of signal events. The loss sensitivity is very small compared to the loss of signal events, so a cut on the Lepton pT 60 GeV is the best cut.

6 Conclusion

After getting the optimizations results we observed that the best cut for the Leading jet pT is between 200 - 300GeV and for the Lepton pT is 60GeV. After this, a final selection will use multivariable analysis. This analysis will be looking into using a deep neural network (Tensorflow), or boosted decision tree (TMVA). In order to analyze all the cuts it is required to use "mini trees". Mini trees are created to store only the relevant information needed to train the machine learn algorithms.

In parallel to the optimization studies, I helped developed the framework with postdoc to create the "mini trees" so the analysis can move into the deep neural network stage.

References

[Col] The ATLAS Collaboration. *Search for pair production of a new heavy quark that decays into a W boson and a light quark in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector.* URL: `https://arxiv.org/pdf/1509.04261.pdf`. (accessed: 05.15.2022).