A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS experiment

Benjamin Warren

PhD student, Department of Physics, University of Cape Town, Rondebosch 7701, South Africa

E-mail: wrrben001@myuct.ac.za

Abstract. A search for tWZ production using events containing three leptons from Run 2 ATLAS proton-proton collision data with a centre of mass energy of 13 TeV will be presented. An event selection scheme was developed using simulation to select tWZ events and to broadly suppress background events. Events were then separated into mutually-exclusive regions of phase space to increase the amount of tWZ events compared to background events, and to calibrate the modelling of the backgrounds. Background events were further suppressed through the use of Gradient Boosted Decision Tree (GBDT) machine learning algorithms. An event-level GBDT was used to distinguish between tWZ and all backgrounds. Using the output score of the event-level GBDT, a maximum likelihood fit, blind to data in the Signal Region, was used to estimate the signal strength, μ, of tWZ production, where nuisance parameters were assigned to theoretical and experimental systematic uncertainties. A signal strength of $\mu = 1.20^{+1.37}_{-1.29}$ was determined with an expected significance of 0.77σ, and an expected upper limit on μ of $2.61^{+2.25}_{-1.21}$ was also determined. These preliminary blinded results show that the search has the potential to put the strongest ever constraint on tWZ production, but does not have the potential to observe tWZ production as predicted by the Standard Model.

1. Introduction

The Standard Model (SM) of particle physics describes the current understanding of the fundamental components of matter and three of the four fundamental forces. It has thus far stood up to every experimental test to which it has been subjected, and with the discovery of the Higgs Boson [1, 2] in 2012, the last major piece of the SM has been confirmed. However, even with the success of the SM, it is unable to address some unanswered questions about matter and forces in the universe. Gravity is not included in the SM, there is no explanation for why the Higgs boson mass seems to arise from very precise cancellations of different contributions [3], and there is no explanation for dark matter [4]. The current goal of particle physics is to search for solutions to these problems by testing the predictions of the SM in the hope of finding hints of Beyond the SM (BSM) physics that could provide solutions to these problems.

The Large Hadron Collider (LHC) is capable of colliding particles at high energies that have never been attained in a lab environment, and is capable of producing collisions at rates far higher than previously achieved. Due to its high energy and high collision rate capabilities, the LHC can produce a significant number of top quarks. This is of interest because as the heaviest elementary particle, the top quark serves as a unique probe in searches for BSM physics at high energies [5]. Furthermore, the ATLAS detector [6] is one of two general purpose experiments at the LHC designed to collect pp collision data that can be used to test the properties of the top quark.

A search is presented for a rare and unobserved production process of a top quark produced in association with a W boson and a Z boson (tWZ production) using run 2 data from the ATLAS detector at the LHC. This is done with the goal of constraining the electroweak couplings of the top quark since these serve as interesting probes of BSM physics [5].The Signal Region (SR), which is the region where the measurement of the cross section is performed, remains blind to run 2 data as this search is part of an ongoing ATLAS analysis and one does not want to be biased by looking at the data in the SR before all other aspects of this analysis are in place.

2. tWZ production and backgrounds in the trilepton channel

In order to reduce the number of background events in the search for tWZ production, a trilepton decay channel was chosen. In Figure 1, an example Feynman diagram of this decay channel is shown. This channel was chosen because in the case of less than three leptons, hadronic background processes dominate to such an extent as to make it impossible to make a measurement of tWZ production in these channels. The trilepton and four lepton channels were considered for further study and the trilepton channel was chosen because it offered manageable background production processes and offered reasonable statistical power.

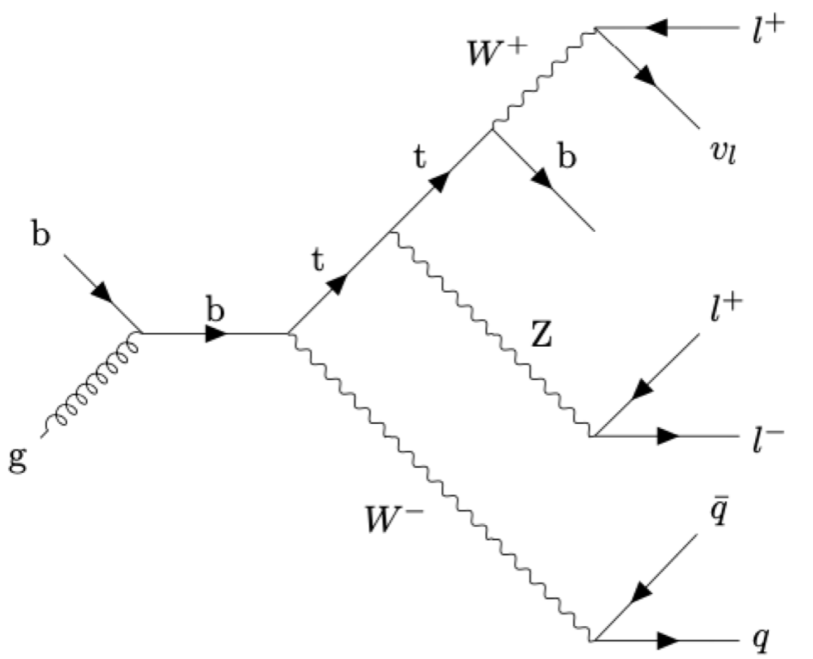

Figure 1. An example Feynman Diagram is shown for tWZ production in the trilepton channel.

In the trilepton search, the WZ and $t\bar{t}Z$ production processes are the dominant source of background events. In order to reduce these backgrounds, as well as other smaller backgrounds, kinematic selections and Machine Learning algorithms were used. This is necessary because of the similar decay products being produced by these processes in the trilepton channel, as well as the much larger cross sections of these processes when compared to tWZ production.

Before trying to distinguish between these two processes and tWZ production, one must check that these processes are well modelled. This is done using a Control Region (CR) for each of these main process. The CRs are also used to constrain these backgrounds in the fit. In Figure 2, the SR for tWZ production and the CRs for WZ and $t\bar{t}Z$ are shown. The simulation for the dominant backgrounds and smaller backgrounds are shown by stacked coloured histograms, and the Run 2 data is shown by the black dots in the two CRs. A normalized distribution of tWZ events is shown by a dotted black line and this shows that tWZ production is most abundant in the SR as desired. The simulation and data agree within uncertainty. This implies that the simulation is well modelled by the data and because these regions are enhanced in the main backgrounds, this implies that these main backgrounds are well modelled by simulation.

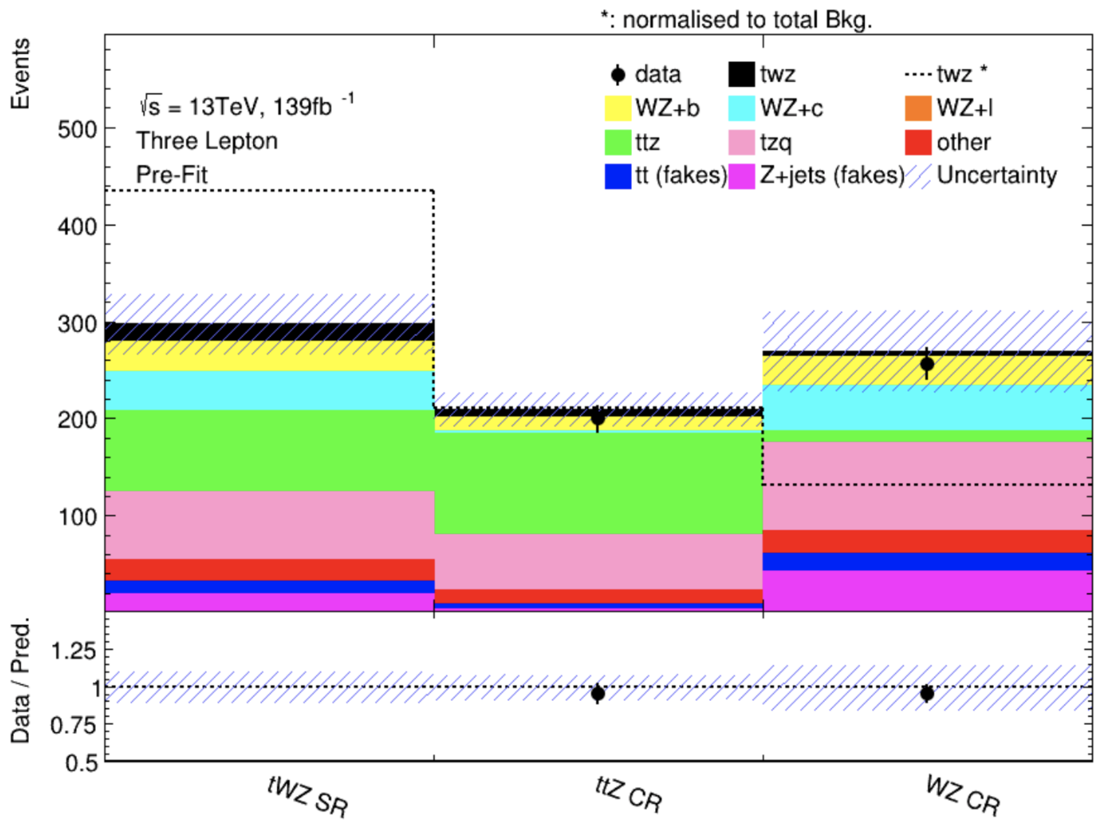

Figure 2. The tWZ SR, WZ CR, and the $t\bar{t}Z$ CR are shown. The SR is blind to data and Run 2 data is shown in the CRs.

3. Event-level GBDT

In this search, an event-level Gradient Boosted Decision Tree (GBDT) is used to distinguish between tWZ production and its main backgrounds. A GBDT is a machine learning algorithm that is trained on several input features to learn the differences between tWZ and the backgrounds. The chosen input features contain information that is useful for making this distinction. In this case, the input features were kinematics of objects in the event, and high level variables that made use of key differences between tWZ production and the main backgrounds.

 ISBN: 978-0-620-97693-0

After training on the input features, the event-level GBDT gives output scores for events based on the values of the features. A score closer to 0 is deemed to be more like a background event, and a score closer to 1 deemed to be more like a signal event. In Figure 3, the output for simulation in the tWZ SR is shown, with the normalized number of signal events shown by a dotted line. According to this figure, the event-level GBDT is effectively distinguishing between tWZ production and the main backgrounds since the dotted line shows more events with higher scores.

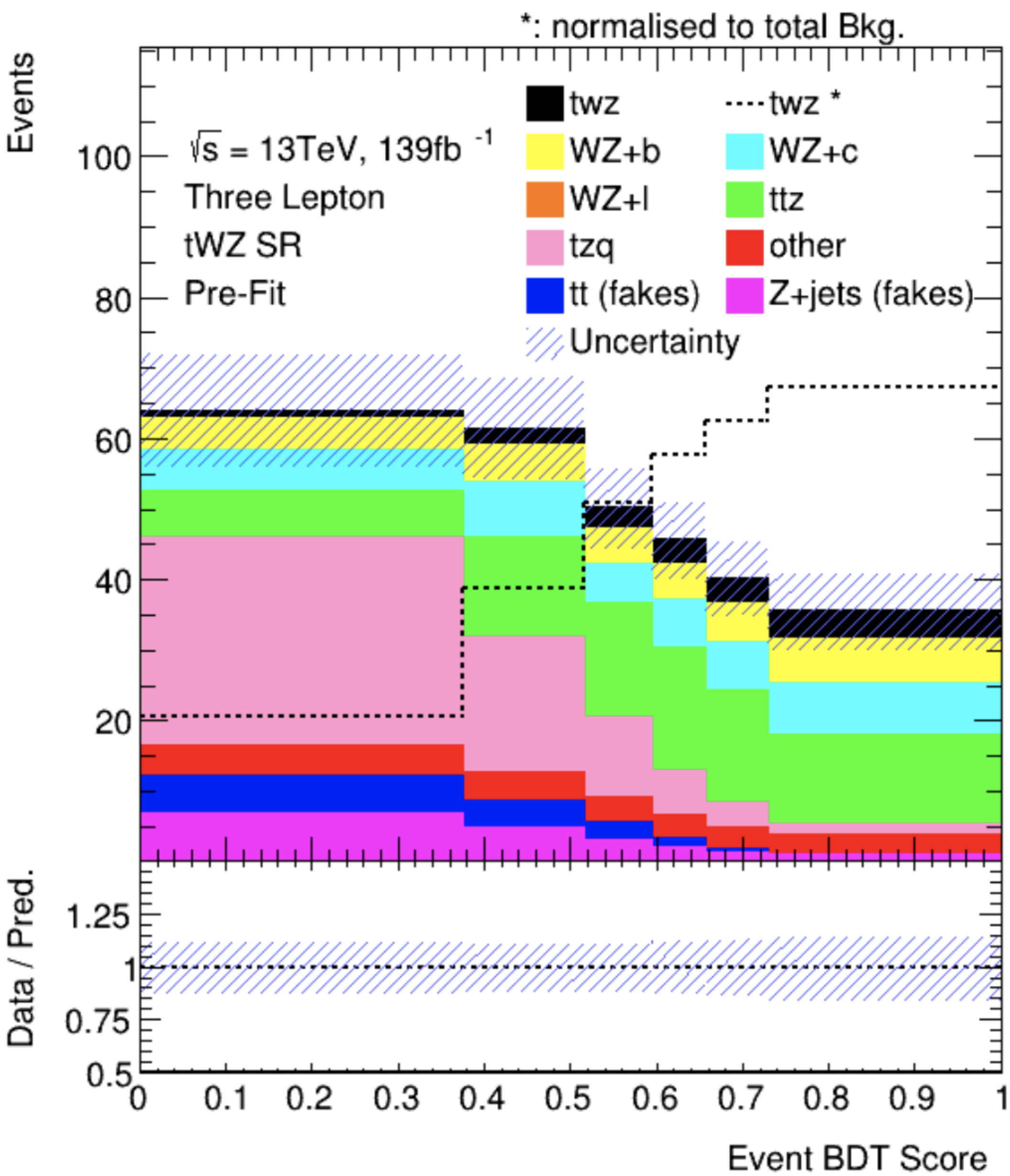

Figure 3. A histogram of the event-level GBDT scores is shown for simulation. A normalized distribution of tWZ events is shown by a dotted line.

 ISBN: 978-0-620-97693-0

4. Results

The signal strength, μ, of a process is the observed cross section divided by the SM prediction of the cross section (σ_{obs}/σ_{SM}). This is used easily see any deviations from the SM prediction. The signal strength of tWZ production is determined using the TRexFitter framework [7] for binned template profile likelihood fits. To determine the signal strength, event-level GBDT score histograms are used in the tWZ SR and the CRs. Since this search is blind in the SR, a dataset constructed from a background-only fit in the CR, is used in the SR to give a more accurate estimation of the sensitivity of this search.

Using the TRexFitter framework [7], a signal strength of $\mu = 1.20^{+1.37}_{-1.29}$ is determined with an expected significance of 0.77σ. This value is not at the 3σ level required in particle physics to constitute evidence. Therefore, an expected upper limit is placed an the signal strength as a measure of the sensitivity of the current setup. An expected upper limit of $2.61^{+2.25}_{-1.21}$ was determined.

From this fit, the main systematic uncertainties can be determined. In Figure 4, a ranking plot of the main systematic uncertainties are shown with those having the greatest impact appearing at the top. Based on this figure, the normalization uncertainties on the cross sections of WZ and $t\bar{t}Z$ production still have the greatest impact on this measurement, and a future measurement would greatly benefit from further suppressing these backgrounds. However, this measurement is strongly statistically limited and would benefit more from Run 3 data from the ATLAS detector.

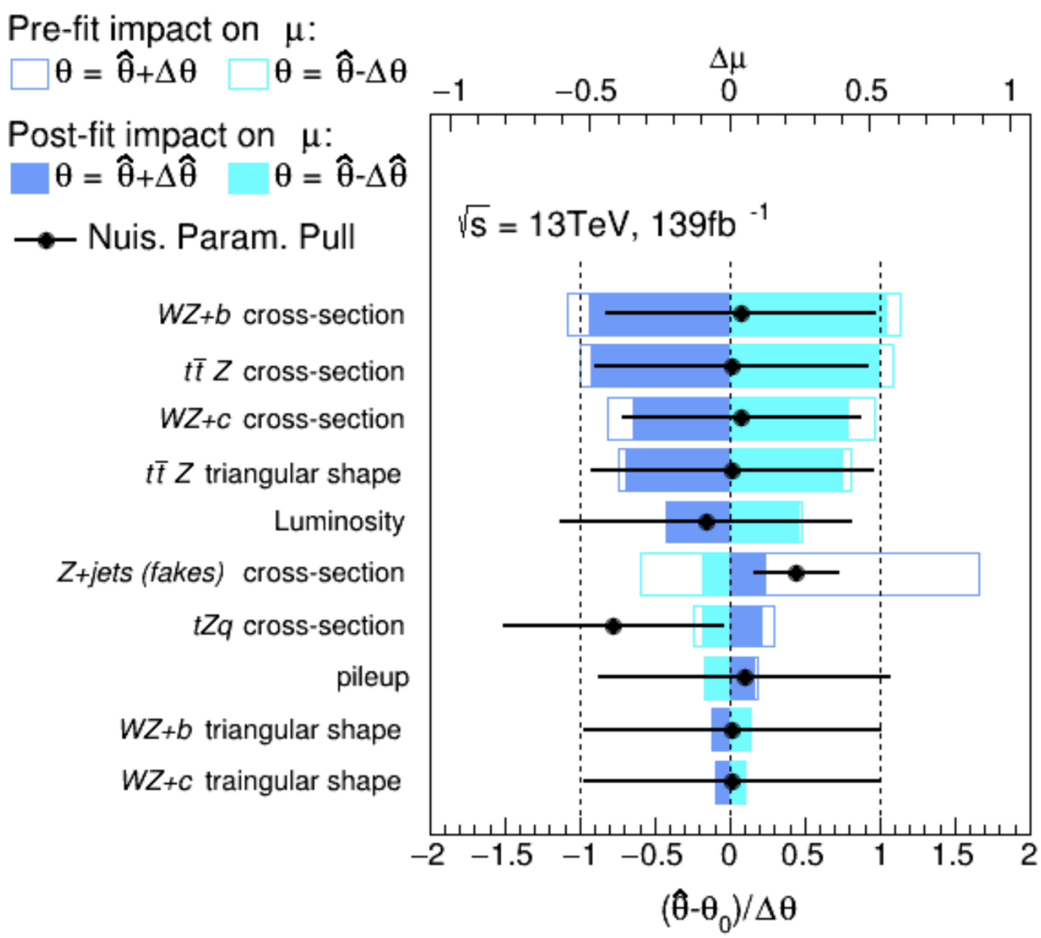

Figure 4. A ranking plot of the main systematic uncertainties ordered by impact is shown.

ISBN: 978-0-620-97693-0

5. Conclusion

In the search for tWZ production in the trilepton channel, data and simulation from the Full Run 2 dataset of the ATLAS detector was used. An event selection scheme was applied to select tWZ events and to broadly suppress background events. Signal and Control Regions were chosen to perform this search and check the modelling of the background processes. Then GBDTs were used to suppress the main backgrounds.

A signal strength of $\mu = 1.20^{+1.37}_{-1.29}$ was determined with an expected significance of 0.77σ, as well as an expected upper limit of $2.61^{+2.25}_{-1.21}$. Given these expected results, one expects to be able to put the tightest limits ever on this rare and interesting production process.

References

[1] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[3] Gian Francesco Giudice. "Naturally Speaking: The Naturalness Criterion and Physics at the LHC". In: *Perspectives on LHC Physics* (2008), 155–178. DOI: 10.1142/9789812779762_0010. URL: http://dx.doi.org/10.1142/9789812779762_0010.

[4] Vera C. Rubin. "One Hundred Years of Rotating Galaxies". In: *Publications of the Astronomical Society of the Pacific* 112.772 (2000), pp. 747–750. DOI: 10.1086/316573. URL: https://doi.org/10.1086/316573.

[5] Fabio Maltoni, Luca Mantani, and Ken Mimasu. "Top-quark electroweak interactions at high energy". In: *Journal of High Energy Physics* 2019.10 (2019). ISSN: 1029-8479. DOI: 10.1007/jhep10(2019)004. URL: http://dx.doi.org/10.1007/JHEP10(2019)004.

[6] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088/1748-0221/3/08/s08003.

[7] Michele Pinamonti and Loic Valery. *TRExFitter*. URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/.