A search for tWZ production at $\sqrt{s}$=13TeV with the ATLAS detector

Alex Veltman, James Keaveney

Department of Physics, University of Cape Town, South Africa, 7701

E-mail: `vltale001@myuct.ac.za`

Abstract. The production of a single top quark in association with a W boson and a Z boson (tWZ) is a rare Standard Model process that has never before been measured. The process is a valuable input into global Standard Model Effective Field Theory (SMEFT) fits due to its sensitivity to top-electroweak SMEFT coefficients in regions of high W boson and Z boson transverse momenta. It is also a relevant background in other top quark related measurements such as the $t\bar{t}Z$ cross-section measurement. A search has been performed for tWZ production using $140\,\mathrm{fb}^{-1}$ of proton-proton collision data at a centre of mass energy of 13 TeV. The search explores the trilepton and tetralepton tWZ final states. Regions are defined using physics object multiplicities and graph neural networks are employed to perform signal-background discrimination in the regions. The signal strength of tWZ production μ_{tWZ} is extracted using a profile likelihood fit with a full systematic model describing experimental and modelling uncertainties. Preliminary blinded measurements of μ_{tWZ} are shown using an Asimov data set for each channel. A combined signal extraction across both channels will also be presented. The measurements of μ_{tWZ} shown will include their associated expected significances and expected upper limits.

1. Introduction

The production of a single top in association with a W boson and a Z boson (tWZ) is a rare process predicted by the Standard Model (SM). The tWZ process has never been observed due to a small expected cross-section of 161 fb and difficult to discriminate sources of backgrounds. An example of a leading order diagram of tWZ production is shown in Figure 1, which includes two electroweak vertices. Due to the presence of top-electroweak vertices, the process is seen as valuable for global Standard Model Effective Field Theory (SMEFT) fits [1]. The rate of tWZ production at high energies is sensitive to certain top-electroweak SMEFT operators, which will allow tWZ production to provide a unique constraint in global SMEFT fits. The process is also a background in other top quark related analyses such as the cross-section measurement of $t\bar{t}Z$ production.

The analysis is divided into the trilepton and the tetralepton decay channels, where both channels require the final state Z boson to decay leptonically. The trilepton decay channel requires either the prompt W boson or the W boson originating from the decaying top quark to decay leptonically. In the trilepton channel, $t\bar{t}Z$ production and the production of a WZ pair with additional jets(WZ+jets) contribute background events that are difficult to distinguish from signal events. The tetralepton channel requires both the prompt W boson and the W boson originating from the decaying top quark to decay leptonically. In the tetralepton channel,

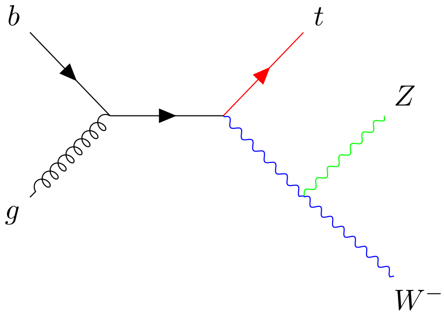

Figure 1. Example of a leading order Feynman diagram for tWZ production. The final state products of a t quark (red), a Z boson (green) and a W boson (blue) are shown.

the major backgrounds are $t\bar{t}Z$ production and the production of a ZZ pair with additional jets (ZZ+jets). The background events in the tetralepton channel are easier to distinguish from signal events in the trilepton channel. The tetralepton channel has fewer total events, which increases the statistical uncertainty associated with the channel compared to the trilepton channel.

2. Samples and event selection

Data and simulated samples were prepared for the trilepton and tetralepton channels. The search was performed using 140 fb^{-1} of 13 TeV proton-proton collisions recorded using the ATLAS detector [2]. This dataset was recorded at the Large Hadron Collider between 2015 and 2018. A comparable dataset of simulated Standard Model processes was generated to compare the event counts between ATLAS data and the Standard Model prediction. The simulated processes include tWZ production as well as the background processes of $t\bar{t}Z$ production and diboson (WZ+jets/ZZ+jets) production with additional jets. Simulated samples were also created for minor multibosonic and top-related background processes.

Leptons, jets and missing transverse momentum were reconstructed from experimental signatures in the various subdetectors of the ATLAS detector. Only electrons and muons were considered. Jets that contain B-hadrons were identified using a machine learning discriminator DL1r [3], where b-tagged jets are defined as jets that pass the 77% DL1r b-jet tagging efficiency operating point.

Events in the trilepton channel were required to have three leptons, where the leading lepton, subleading and subsubleading lepton must have $p_T > 30$GeV, $p_T > 20$GeV and $p_T > 14$GeV, respectively. Muons must be within the region of $|\eta| < 2.5$ and electrons must be within $|\eta| < 2.47$ and not within $1.37 < |\eta| < 1.52$. All jets must have $p_T > 25$ GeV, be in the forward region of $|\eta| < 2.5$ and have a jet-vertex-tagger (JVT) [4] of greater than 0.5. Any combination of oppositely-signed same-flavoured (OSSF) leptons must have a combined mass that is greater than 10 GeV. A combined OSSF lepton pair is a Z candidate if the pair's mass is within 10 GeV of the Z boson mass $m_Z = 91.19$. Events in the trilepton channel were required to have exactly one Z candidate.

Each event in the tetralepton channel must have four leptons, where the highest energy lepton has $p_T > 28$ GeV, the second highest energy lepton has $p_T > 18$ GeV and the third and fourth highest energy leptons have $p_T > 10$ GeV. The jet selection requirements and the lepton η requirements are the same as for the trilepton channel. The sum of the lepton charges must also be neutral and all OSSF lepton pairs in the event must have a mass greater than 10GeV.

All events were partitioned into regions based on the multiplicity. The trilepton channel has three regions: one signal region (SR) and two control regions (CR). The events in the trilepton tWZ SR were required to have three or more jets, where exactly one of the jets is a b-tagged jet. The events in the trilepton $t\bar{t}Z$ CR must have four or more jets, where two or more of the jets are b-tagged jets. The events in the trilepton WZ CR must have one or two jets, where exactly one jet is a b-tagged jet. Five regions were defined for the tetralepton channel. The tetralepton channel has two signal regions that require events to have one Z candidate, greater than one jet and exactly one b-tagged jet. The tWZ opposite-flavoured (OF) SR and the tWZ same-flavoured (SF) SR differ based on the flavours of the leptons that did not originate from the Z candidate. The tWZ OF SR contains events where the two non-Z leptons have different flavours and the tWZ SF SR contains the events where the two non-Z leptons have the same flavour.

3. Background discrimination with graph neural networks

The signal regions are dominated by background contributions from $t\bar{t}Z$ and diboson production. A pair of graph neural networks [5] were developed to discriminate between signal and background events in the trilepton and tetralepton channels. Each event can be modelled as a graph where the muons, electrons, jets and missing transverse momentum are nodes in the graph. Each node carries information regarding the 4-momentum of the object and what type of object the node represents. The edges between the nodes carry an edge weight that is equal to the angular difference ΔR between the objects associated with the nodes. Additional graph-level information was also included regarding the number of jets in the event. The graph models followed the message-passing neural network architecture [6].

The model used in the trilepton channel was trained and tested using simulated tWZ, $t\bar{t}Z$ and WZ+jets events from the trilepton tWZ SR. The model used in the tetralepton channel was trained and tested using simulated tWZ, $t\bar{t}Z$ and ZZ+jets events from the tetralepton signal regions and the $t\bar{t}Z$ CR. The models are binary classifiers that produce a score between 0 and 1, where the signal tWZ events were assigned a label of 1 and the background $t\bar{t}Z$ and diboson events were assigned a label of 0.

4. Systematic uncertainties

The search considers the experimental uncertainties related to data-taking and the modelling uncertainties associated with the various simulated processes. The recommendations regarding systematic uncertainties provided by the internal ATLAS combined performance groups were followed. The experimental uncertainties associated with the luminosity and pileup conditions of the datasets were considered. A set of uncertainties related to the reconstruction of objects were included in the analysis, such as uncertainties associated with the energy scale and resolution of objects, jet flavour information and lepton efficiency scale factors.

The modelling uncertainties in the search included the estimation of the cross section normalisation factor applied to each of the simulated background samples. The choice of diagram removal scheme used for the tWZ sample [7] was accounted for using a systematic uncertainty. For the tWZ sample, systematic uncertainties were introduced to describe normalisation and factorisation scale variations in the matrix element calculations and to describe different parameterisations of the parton distribution function (PDF).

5. Extraction method

The parameter of interest of the extraction method is the signal strength of tWZ production, $\mu_{tWZ} = \sigma_{\text{obs}}/\sigma_{SM}$, where σ_{obs} is the observed cross-section of tWZ production and σ_{SM} is the expected cross-section of tWZ production according to the SM. A likelihood model [8] was constructed with μ_{tWZ} as the parameter of interest as well as nuisance parameters describing each systematic uncertainty. The parameter values were estimated using a profile likelihood fit [9], which allows for the statistical significance Z_0 of the measurement to be obtained.

In this proceeding, only an *Asimov dataset* [9] was used during the signal extraction. An Asimov dataset is designed to produce best-fit values of the parameters that are exactly equal to the simulated expectation. The Asimov dataset allows for the experimental sensitivity of the analysis strategy to be evaluated. A *stat-only fit* was performed to evaluate the uncertainty due to statistical fluctuations. This involves first performing a fit to data with the parameter of interest fixed in order to estimate the other parameters. The parameter of interest is then estimated while the other parameters are kept fixed at these new values.

6. Fit results with Asimov dataset

The extraction of the tWZ signal strength μ_{tWZ} was performed individually on the trilepton and tetralepton Asimov datasets and a combined extraction was performed using both Asimov datasets. The measured signal strength μ_{tWZ} was obtained through profile likelihood fits using the trilepton dataset, tetralepton dataset and combined dataset and the results are shown in Figure 2. All μ_{tWZ} were found to be equal to one, which is expected when performing fits to an Asimov dataset.

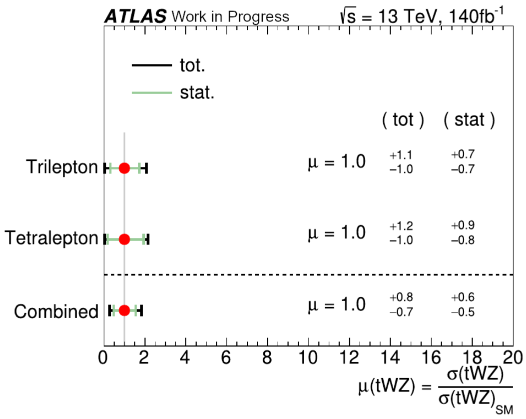

Figure 2. The estimated value and uncertainty of μ_{tWZ} for various fits to different Asimov datasets. This includes the fit results from the trilepton-only, tetralepton-only and combined fits. The total uncertainty of the fit is shown. The uncertainty when a stat-only fit is performed is also shown.

The total uncertainty associated with the measurements in the trilepton channel and the tetralepton channel were similar. The precision of the trilepton measurement was affected by both systematic and statistical uncertainty. The measurement in the tetralepton channel was dominated by statistical uncertainty. The relatively high statistical uncertainty was attributed to the difficulty in separating tWZ events and background events as well as the low number of events with four leptons. The combined measurement has a lower total uncertainty compared to the measurements in the trilepton or tetralepton channels alone.

The post-fit values of the nuisance parameters describing the systematic uncertainties are ranked in terms of impact in Figure 3. The estimation of the cross-section normalisation of $t\bar{t}Z$ production had the largest impact on the measurement, which may be the result of $t\bar{t}Z$ production being a background in both the trilepton and tetralepton channel. The systematic uncertainty associated with the choice of event generator used in simulating the $t\bar{t}Z$ sample is also highly ranked in impact. The uncertainties associated with jet energy resolution and the variation in the PDF calculations for the tWZ samples were the second and third most impactful systematic variables.

The expected significance associated with the μ_{tWZ} measurement in the combined fit was $Z_0^{\text{exp}} = 1.34\sigma$. The expected significance of the measurement suggests that a discovery of tWZ production will not be possible if the observed ATLAS data agrees with the simulated samples. A reduction in the systematic and statistical uncertainty would be required for an observation to be expected.

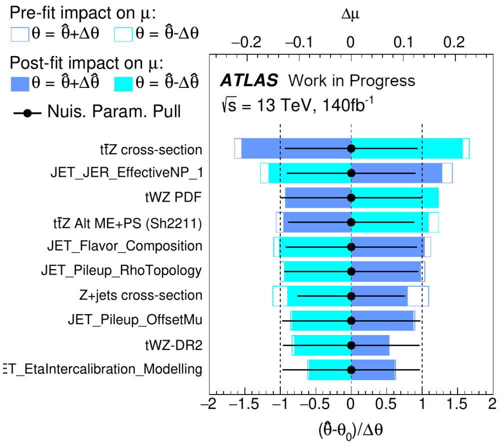

Figure 3. Ranking of pre-fit and post-fit impact values of each nuisance parameter. Only the nuisance parameters with the ten highest impact values are shown. The upper axis is in terms of impact $\Delta\mu$. The pre-fit impacts are shown as unfilled blue rectangles and the post-fit impacts are shown as filled blue rectangles. The lower axis describes pull on the nuisance parameter during the fit, where $\hat{\theta}$, θ_0 and $\Delta\theta$ are the best-fit value, pre-fit value and pre-fit uncertainty of the nuisance parameter. The pull value and post-fit uncertainty of each nuisance parameter are included as black points.

7. Conclusion

The production of a single top quark in association with a W boson and a Z boson (tWZ) is a rare Standard Model process that has never before been measured. A search for tWZ production was performed in the trilepton and tetralepton final states using $140\,\text{fb}^{-1}$ of proton-proton collision data taken at the ATLAS detector. The experimental sensitivity of the analysis strategy was demonstrated using Asimov datasets. A signal strength value of $\mu_{tWZ} = 1.0^{+0.8}_{-0.7}$ was found when using a combined trilepton and tetralepton Asimov dataset. The associated expected significance of the measurement was $Z_0^{\text{exp}} = 1.34\sigma$. The measurement was found to be both statistically and systematically limited.

References

[1] Maltoni F, Mantani L and Mimasu K 2019 *JHEP* **10** 004 (*Preprint* `1904.05637`)

[2] Collaboration T A 2008 *Journal of Instrumentation* **3** S08003 URL `https://dx.doi.org/10.1088/1748-0221/3/08/S08003`

[3] ATLAS Collaboration 2019 *The European Physical Journal C* **79** URL `https://doi.org/10.1140/epjc/s10052-019-7450-8`

[4] 2014 Tagging and suppression of pileup jets Tech. rep. CERN Geneva URL `https://cds.cern.ch/record/1643929`

[5] Battaglia P W, Hamrick J B, Bapst V, Sanchez-Gonzalez A, Zambaldi V, Malinowski M, Tacchetti A, Raposo D, Santoro A, Faulkner R, Gulcehre C, Song F, Ballard A, Gilmer J, Dahl G, Vaswani A, Allen K, Nash C, Langston V, Dyer C, Heess N, Wierstra D, Kohli P, Botvinick M, Vinyals O, Li Y and Pascanu R 2018 Relational inductive biases, deep learning, and graph networks (*Preprint* `1806.01261`)

[6] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E 2017 Neural message passing for quantum chemistry (*Preprint* `1704.01212`)

[7] Bessidskaia Bylund O (ATLAS) 2016 Modelling Wt and tWZ production at NLO for ATLAS analyses *9th International Workshop on Top Quark Physics* (*Preprint* `1612.00440`)

[8] Cranmer K, Lewis G, Moneta L, Shibata A and Verkerke W (ROOT) 2012 HistFactory: A tool for creating statistical models for use with RooFit and RooStats Tech. rep. New York U. New York URL `https://cds.cern.ch/record/1456844`

[9] Cowan G, Cranmer K, Gross E and Vitells O 2011 *The European Physical Journal C* **71** URL `https://doi.org/10.1140/epjc/s10052-011-1554-0`