IL NUOVO CIMENTO **47 C** (2024) 131
DOI 10.1393/ncc/i2024-24131-y

Colloquia: IFAE 2023

First t-channel single top measurement in pp collisions at an energy in the centre-of-mass of $\sqrt{s} = 5.02$ TeV with the ATLAS experiment(*)

L. Pintucci([1])([2])(**) on behalf of the ATLAS Collaboration

([1]) *INFN Sezione di Trieste, Gruppo Collegato di Udine - Udine, Italy*
([2]) *Università degli Studi di Trieste - Trieste, Italy*

received 13 February 2024

Summary. — A single top quark production measurement in proton-proton (pp) collisions at the Large Hadron Collider (LHC) at an energy in the centre-of-mass $\sqrt{s} = 5.02$ TeV with data collected by the ATLAS experiment is reported. This analysis uses a Boosted Decision Tree (BDT) to separate signal from background events and its output distribution is used for a Profile-Likelihood fit. Some of the calibrations used in this analysis are taken from a recent ATLAS top quark pair production cross-section measurement at the same energy of $\sqrt{s} = 5.02$ TeV (ATLAS Collaboration, *JHEP*, **06** (2023) 138). The analysis takes into account forward jets to enhance the t-channel process, with an ad-hoc calibration for their energy.

1. – The Motivation

Single top production is predicted by the Standard Model (SM) via three main modes: t-channel, s-channel and W boson associated production (also referred to as tW mode). The t-channel is the process with the highest cross-section, while the others are quite less abundant. The expected SM cross-section at $\sqrt{s} = 5.02$ TeV is $\sigma_{SM} = 30.3$ pb. The ATLAS collaboration measured the single top t-channel production already at $7, 8$, and 13 TeV as shown in fig. 1.

In this analysis, the t-channel single top production cross-section is measured with data collected at the centre-of-mass energy of $\sqrt{s} = 5.02$ TeV by the ATLAS experiment [1].

(*) IFAE 2023 - "Poster" session
(**) E-mail: `laura.pintucci@cern.ch`

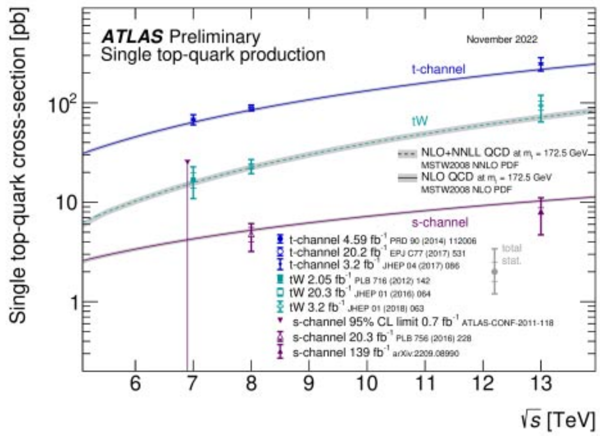

Fig. 1. – Cross-section prediction for t-channel, tW and s-channel single top production, compared with various ATLAS measurements for these processes [2].

2. – Analysis Strategy

2˙1. *Data and Monte Carlo samples.* – The data sample used was collected in 2017 with an integrated luminosity of $L = 257$ pb^{-1} at the centre-of-mass energy of $\sqrt{s} = 5.02$ TeV and an average number of interactions per pp bunch crossing which ranged from approximately 0.5 to 4 with an average of 2 interactions.

Monte Carlo (MC) samples were generated to be compared with data, they are processed through the full ATLAS detector simulation in Geant4 [3]. The t-channel signal process is simulated with Powheg [4] and Pythia [5]. The same Monte Carlo programs are also used for the background s-channel, tW mode and top quark pair ($t\bar{t}$) production samples, while background samples for W or Z boson associated with jets and diboson production are simulated with Sherpa [6].

2˙2. *Event Selection.* – The t-channel single top production comes from the exchange of a W boson between a light and a heavy quark as shown in fig. 2.

The top quark decays almost always to a bottom quark (b) and a W boson, which can then decay leptonically into an electron/muon/τ and its neutrino or hadronically in quark-antiquark pairs. This analysis considers the leptonic decay of the W boson.

In order to have events enriched in the signal process, events are selected if they have exactly one lepton (electron or muon) with transverse momentum $p_T > 18$ GeV,

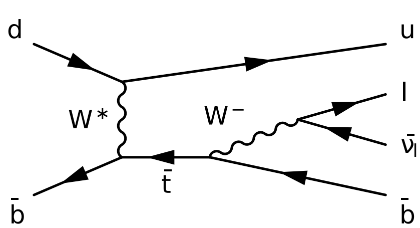

Fig. 2. – Single top t-channel production Feynman diagram

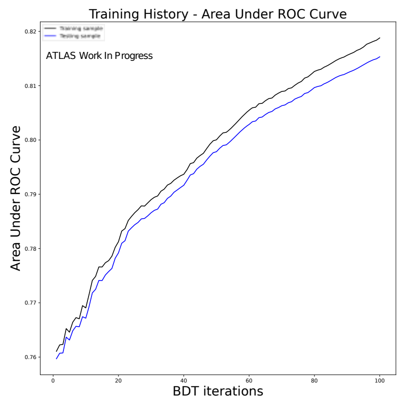

Fig. 3. – Area under the ROC curve from the first to the 100th iteration. The blue line is relative to the validation sample, while the black one to the training sample.

and exactly 2 jets with $p_T > 23$ GeV. Exactly one of these jets must be b-tagged with the DL1r algorithm [7] and it must have a pseudo-rapidity $|\eta| < 2.5$. The topology of the single top t-channel production process is characterized by the light (non b-tagged) jet produced at high pseudo-rapidity, in the forward region of the detector. Jets in the forward region are not usually considered in ATLAS physics analysis, anyway in this work the reconstructed non b-tagged jets with $|\eta| < 4.0$ are used, as this enhances the signal over background ratio. An ad-hoc calibration for the jet energy scale has been used with in-situ correction computed with the sample of data described in sect. **2**˙1.

2˙3. *Boosted Decision Tree.* – In order to enhance the separation of signal and background events a boosted decision tree (BDT) is used.

The BDT is trained on the MC samples for signal and background, and its parameters are optimized to obtain a larger separation between signal and background in the BDT output distribution. The BDT has 9 input features that describe the object kinematics and the global event topology. The validation step is performed with a stratified k-fold with 3 folds. One of the metrics used to validate the performance of the BDT and to check not to have overtraining effects is the area under the ROC curve, which is shown in fig. 3.

2˙4. *Data and Monte Carlo comparison.* – The distributions of various observables describing the kinematic of the reconstructed events are used to compare the Monte Carlo prediction with data, as well as the distribution of the BDT output, as shown in fig. 4.

3. – Asimov Fit and Results

The t-channel single top production cross-section is extracted with a Profile Likelihood fit of the BDT output distribution.

A profile likelihood fit is first performed on an Asimov dataset, a dataset build so that the fit free parameters match the expected values. The Asimov fit is used to test

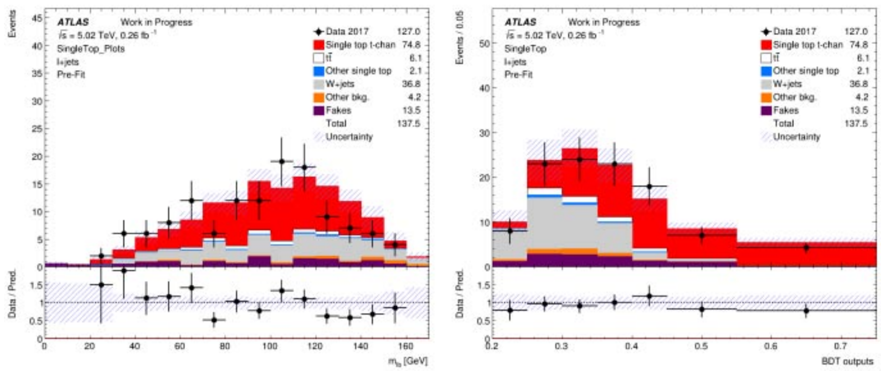

Fig. 4. – Monte Carlo prediction (filled histogram) and data distribution (black dots) on the left for the invariant mass of the lepton and b-tagged jet of the candidate event and on the right for the BDT output.

the full fit setup with the Asimov pseudo-data. Figure 5 shows the nuisance parameters of the Asimov profile likelihood fit ordered based on the impact they have on the final extracted systematic uncertainty. The output of the Profile Likelihood fit is the signal strength μ, which is defined as the ratio between the measured cross-section and the one predicted by the Standard Model.

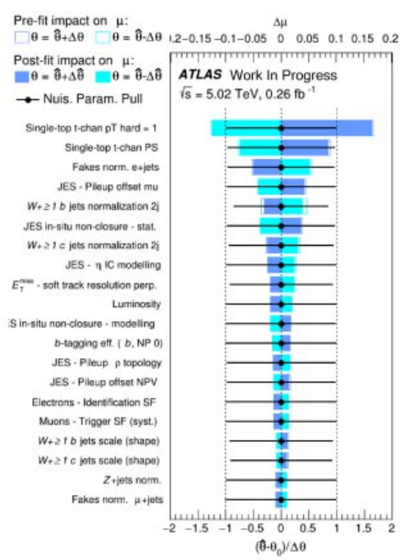

Fig. 5. – The 20 most relevant nuisance parameters chosen and ordered based on the impact they have on the final systematic uncertainty of the Asimov fit.

3˙1. *Results*. – The Asimov fit measures on the pseudo-data the signal strength parameter, which can then be used to obtain the t-channel cross-section. The expected signal strength is found to be $\mu = 1.0 \pm 0.22$, and the background-only hypothesis is rejected with an expected significance of 7.9 standard deviations.

4. – Conclusions

In this report the analysis strategy to extract the t-channel single top production is shown. Kinematics selection on the event topology are used, together with a BDT to separate signal and background events. Good agreement is found between data distributions and Monte Carlo predictions for various kinematic observables describing the event topology. The profile likelihood fit on the Asimov dataset shows that the background-only hypothesis is rejected with an expected significance of 7.9 standard deviations.

REFERENCES

[1] ATLAS Collaboration, *JINST*, **3** (2008) S08003.
[2] ATLAS Collaboration, ATL-PHYS-PUB-2023-014, `http://cds.cern.ch/record/2872786`.
[3] Agostinelli S. *et al.*, *Nucl. Instrum. Methods A*, **506** (2003) 250.
[4] Frederix R., Re E. and Torrielli P., *JHEP*, **09** (2012) 130.
[5] Sjostrand T., Mrenna S. and Skands P. Z., *Comput. Phys. Commun.*, **178** (2008) 852.
[6] Bothmann E. *et al.*, *SciPost Phys.*, **7** (2019) 034.
[7] ATLAS Collaboration, *Eur. Phys. J. C*, **83** (2023) 681.