PROCEEDINGS OF SCIENCE

Evidence for four-top-quarks production with the ATLAS detector at the Large Hadron Collider

Paolo Sabatini[a,*], **on behalf of the ATLAS Collaboration**

[a] *Institut de Física Corpuscular (IFIC), Centro Mixto Universitat de Valencia - CSIC*
Parque Científico, C/Catedrático José Beltrán 2
Paterna, Spain

E-mail: paolo.sabatini@cern.ch

The outstanding performance of the LHC during Run 2 data taking let the ATLAS Experiment to collect a massive dataset of pp collision at $\sqrt{s}$ = 13 TeV. This impressive availability of data to analyse paves the way to the research of rare events, as test of the Standard Model validity. One of the most interesting process is the production of $t\bar{t}t\bar{t}$, an ultra rare process predicted by the Standard Model and very sensitive to new physics contribution. The ATLAS experiment has targeted the measurement of this process in two separate analyses, focused in two different detection channels. These analyses led to the first evidence of $t\bar{t}t\bar{t}$ production with an observed (expected) significance of 4.7 (2.6) standard deviations in their combination.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

*Speaker

PoS(EPS-HEP2021)489

1. Introduction

Thanks to the exceptional luminosity delivered by the LHC [1] during the Run 2, the ATLAS Experiment [2] has collected an extremely large dataset of pp collisions at $\sqrt{s}$ = 13 TeV. Therefore, a deep investigation of the features of the Standard Model (SM) is now possible via the measurement of ultra rare processes, that are predicted by the SM, but not yet observed. One example is the production of the $t\bar{t}t\bar{t}$.

The $t\bar{t}t\bar{t}$ production is one of the rarest processes in the top-quark sector of the SM that is accessible with the current collected luminosity. A representative Feynman diagram is shown in Figure 1a. Its cross-section is predicted by the SM to be 12 fb with an uncertainty of 20%. The calculation is at the Next-to-Leading-Order (NLO) precision including strong and electroweak corrections [3]. Therefore, its measurement represents a test of the SM validity in the top-sector down to fb level.

Indeed, the $t\bar{t}t\bar{t}$ cross-section is also very likely to be altered by contributions of new physics. Heavy Higgs from the 2HDM model may be produced in association with a top pair and then decay in a top pair, enhancing the $t\bar{t}t\bar{t}$ cross-section [4, 5]. Heavy scalar gluons are expected to be produced in pairs at LHC, providing a $t\bar{t}t\bar{t}$ final state by decaying into top-quarks [6, 7]. Otherwise, $t\bar{t}t\bar{t}$ final state can be produced by decay chains of heavy supersymmetric particles [8, 9]. In the context of an effective field theory, the production of four top-quarks is regulated via a contact interaction term, that can be constrained by a measurement of this process. Finally, the $t\bar{t}t\bar{t}$ measurement gives an orthogonal probe of the top-Higgs coupling with respect to the $t\bar{t}H$ production [11, 12].

Experimentally, the measurement of the $t\bar{t}t\bar{t}$ production involves many typologies of final states in terms of lepton multiplicity, summarized in Figure 1b. These final states are grouped into two analysis channels, based on the dominant backgrounds: the same-sign dilepton and multilepton channel (SSML) and the single lepton and opposite-sign dilepton channel (1L/2LOS). The analysis addressing the $t\bar{t}t\bar{t}$ signature in the SSML channel [13] is described in Section 2. The description of the analysis targeting the $t\bar{t}t\bar{t}$ signature in the 1L/2LOS channel [15] is contained in Section 3. Section 3 also contains the results of the combination of the two channels.

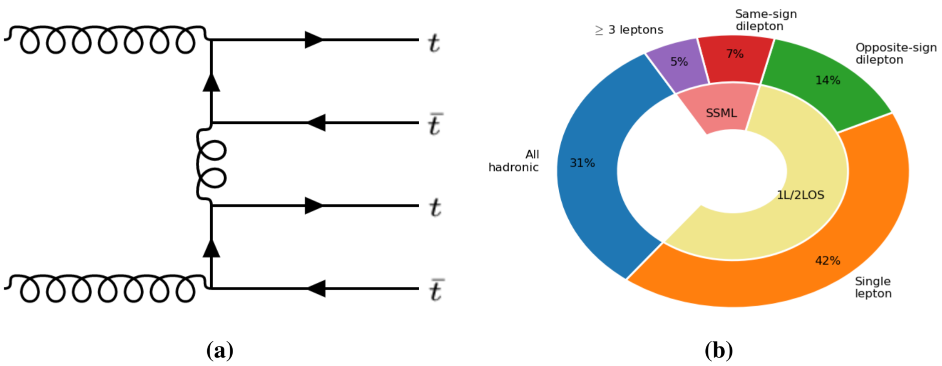


Figure 1: Representative Feynman diagram for the $t\bar{t}t\bar{t}$ production at LHC (a) [15]. Pie chart of the $t\bar{t}t\bar{t}$ final states (b).

2. Same-sign dilepton and multilepton channel

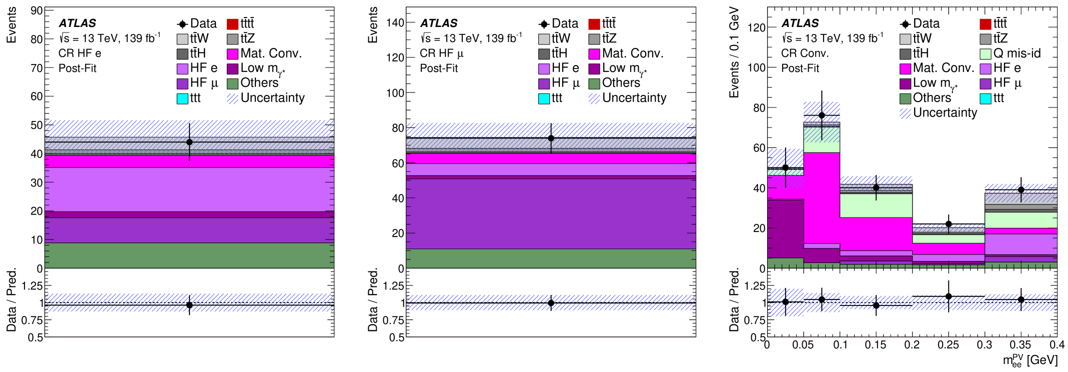
In this section, the analysis performed by the ATLAS Collaboration targeting the $t\bar{t}t\bar{t}$ measurement in the SSML channel is described [13]. The analysis uses the dataset of 139 fb^{-1} collected by the ATLAS experiment during the LHC Run 2.

In this channel, events with two leptons with same-sign charge or at least three leptons are selected. Under these conditions, $t\bar{t}t\bar{t}$ provides still many jets in the final state (at least six), of which many are expected to be induced by b-quarks (so called b-jets). Therefore, at least six jets are required in the selected events. Furthermore, two of the selected jets are required to be tagged as b-jets by the b-tagging algorithm. The large lepton multiplicity in the final state induces a significant contribution of events with non-prompt or fake leptons. As this contribution increases the lower the p_{T} of the objects is, an additional cut of $H_{\mathrm{T}} > 500$ GeV is imposed, where H_{T} is defined as the sum of the p_{T} of all the leptons and jets in the final state.

These selections define the $t\bar{t}t\bar{t}$ signal region: the main sources of backgrounds are given by $t\bar{t}W$ irreducible background (37%) and instrumental background (25%). The estimation of these two sources of background is therefore crucial for a robust measurement of the $t\bar{t}t\bar{t}$ process. Instrumental background are divided in many classes, depending on the corresponding sources. The dominant sources are: non-prompt leptons from heavy-flavour hadron decays, electron pairs from photon conversion or Drell-Yann decays and charge mis-identification. The background coming from charge mis-identification is made of events with pairs of opposite-sign leptons that are reconstructed as same-sign pairs. This background is estimated by measuring the charge mis-identification probability. This probability is estimated with events events that reconstruct the Z-boson resonance by the ratio of yields of events having same-sign pairs of leptons over the opposite-sign dilepton events. The computed values are used to weight events in an opposite-sign dilepton region to the same-sign one. The other backgrounds are instead estimated via template method: their normalisations are measured in dedicated control regions, assuming the shapes of Monte Carlo (MC) simulations. The post-fit agreement of data with the expected background in the control regions are shown in Figure 2. Similarly, the irreducible $t\bar{t}W$ background is estimated via template method, by measuring its normalisation in a dedicated control regions assuming the shape of the distributions from the MC simulation. The $t\bar{t}W$ modeling is further controlled in a validation region, built by loosening the jet multiplicity requirement to four and looking at the difference of yields of events with positive and negative sum of lepton charges. Thanks to the charge asymmetry of the $t\bar{t}W$ production, this region is particularly pure in this source of background. While the control regions are included in the profile likelihood fit together with the signal region, the validation region is only used to validate the background modeling. The excellent post-fit agreement in the $t\bar{t}W$ control and validation regions are shown in Figure 3. The slight excess of data with respect to the expected $t\bar{t}W$ contribution at large jet multiplicity is included in the fit as a additional uncertainties of 130% and 300% on the modeling of $t\bar{t}W$ in association with 7 and $\geq$ 8 jets, respectively. Similarly, an additional uncertainty of 50% on $t\bar{t}W$ production in association with $\geq$ 3 b-jets is included in the fit.

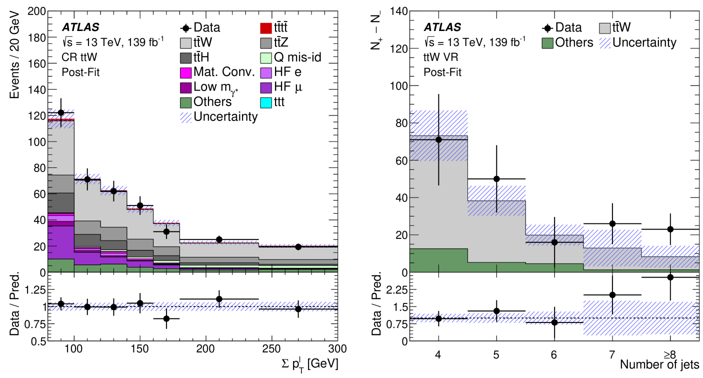
A profile likelihood fit is performed in the signal and control regions simultaneously. The discriminant variable in the signal region consists of the output of a Boosted Decision Tree (BDT), trained in order to maximise the separation between $t\bar{t}t\bar{t}$ and the total background. The signal

PoS(EPS-HEP2021)489



(a) **(b)** **(c)**

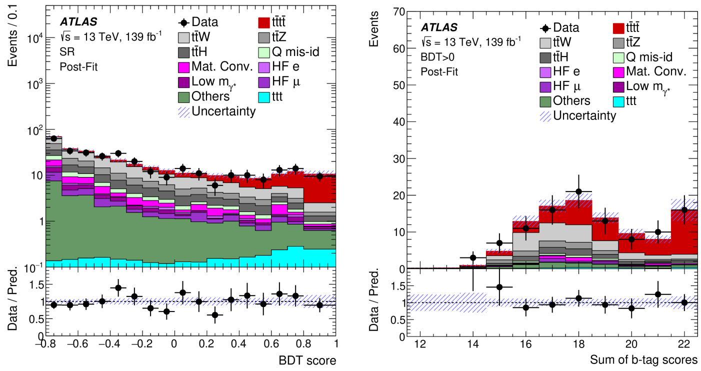
Figure 2: Post-fit agreement of data with the estimated background in the control regions addressing the major sources of instrumental background: non-prompt electrons (a) and muons (b) and material or Drell-Yan conversions (c) [13].



(a) **(b)**

Figure 3: Post-fit agreement of data with the estimated background in the control (a) and validation (b) regions addressing the $t\bar{t}W$ background [13].

strength of the $t\bar{t}t\bar{t}$ production is fitted together with the normalisation factors for all the main backgrounds and the nuisance parameters associated to the sources of instrumental and modeling systematic uncertainties. The normalisation factors for the background are fitted to be consistent with unity within the uncertainties. The only exception is the normalisation of $t\bar{t}W$, estimated to be 1.6 ± 0.3, consistently with previous ATLAS results in a similar phase space [14]. The post-fit distribution of the BDT output in the signal region is shown in Figure 4a, corresponding to a cross-section of 24^{+7}_{-6} fb for the $t\bar{t}t\bar{t}$, compatible with SM expectation within two standard deviations. A pure $t\bar{t}t\bar{t}$ sample is selected by requiring the BDT output to be positive in the signal region: the excess of data with respect to total background is well compatible with the $t\bar{t}t\bar{t}$ hypothesis, as shown in Figure 4b. The observed (expected) significance of the $t\bar{t}t\bar{t}$ signal over the background-only hypothesis is 4.3 (2.4) standard deviations, providing the first evidence of this process.



(a) **(b)**

Figure 4: Post-fit agreement of data with the estimated background in the signal region (a). Distribution of the sum of the b-tagging algorithm scores of all the jets in the signal region events in a $t\bar{t}t\bar{t}$-enhanced region, obtained by requiring a positive BDT score (b) [13].

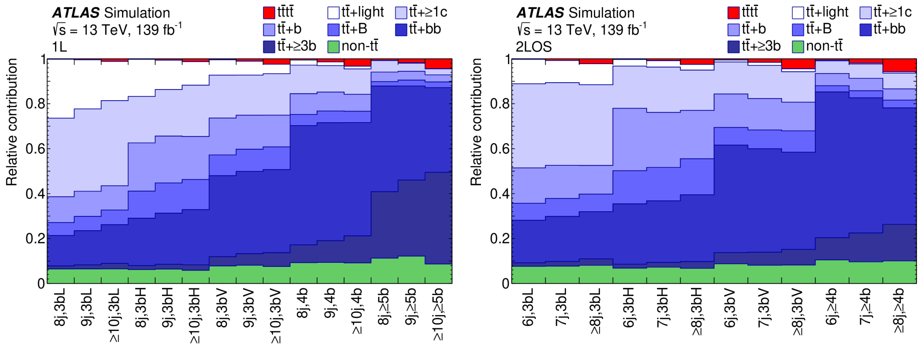
3. Single lepton and opposite-sign dilepton channel

In this section, the analysis performed by the ATLAS Collaboration targeting the $t\bar{t}t\bar{t}$ measurement in the 1L/2LOS channel is described [15]. Similarly to the analysis in the SSML channel, this analysis uses the dataset of 139 fb^{-1} collected by the ATLAS experiment during the LHC Run 2.

This analysis selects events with one lepton or two leptons with opposite-signed charges in the final state. Differently from the analysis in the SSML, that deals with many sources of background, this channel has only a single dominant source of background: $t\bar{t}$ production associated with a large jet radiation. As the background has two top-quarks less in the final state, six jets (of which two are b-jets) are required to be radiated by the $t\bar{t}$ background to cover the same phase-space as the signal. A classification of the $t\bar{t}$ events in terms of the flavour of the radiation is therefore needed to control the type of events that enter the signal region. This classification is based on the truth information of the $t\bar{t}$ simulation, by matching the reconstructed jets with simulated heavy-flavour c- or b- hadrons. The $t\bar{t}$ production is divided into $t\bar{t}$+≥1b, $t\bar{t}$+≥1c and $t\bar{t}$+light jets. Furthermore, the $t\bar{t}$+≥1b is splitted in bb, B and ≥ 3b based on the number of b-hadrons that are matched to the reconstructed jets[1]. In order to control the evolution of the different compositions of the $t\bar{t}$ radiation, events with at least 8 (6) jets with at least 2 b-jets are selected in the 1L (2LOS) channel and classified based on jet and b-jet multiplicities. The composition of the background in the analysis regions is shown in Figure 5. The regions requiring three b-jets in the final state, labeled with 3b, are furthermore split in 3bL, 3bH and 3bV. These regions are defined by using different working points of the b-tagging algorithm for the third b-jet definition. The 3bL uses the loosest working point, providing a larger contribution of $t\bar{t}$+≥1c, while the 3bV uses the tightest working point providing a larger contribution of of $t\bar{t}$+≥1b, closer to the signal region.

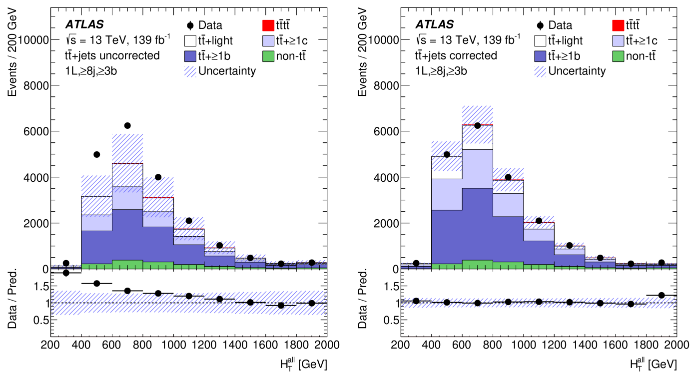
The simulation of the $t\bar{t}$ process is used to provide the background estimation in the analysis regions. However, an evident mismodeling of this process in this extreme phase space is observed, as shown in Figure 6a. Therefore, a MC-based background estimation procedure is performed to

[1]The label B corresponds to a reconstructed jet that is matched with two simulated b-hadrons.



(a) (b)

Figure 5: Composition of the events in the analysis region in the single lepton (a) and opposite-sign dilepton (b) channels. The selection applied on each region is encoded in the region naming as: $XjYb$, where X jets and Y b-jets are required [15].



(a) (b)

Figure 6: Agreement of the data with expected $t\bar{t}$ background in the analysis region before (a) and after (b) the background estimation procedure. Here, an inclusive region requiring at least 8 jets and 2 b-jets in the single lepton channel is shown [15].

improve the background modeling. The first step of this procedure consists in the estimation the normalisation factors for the $t\bar{t}$ radiation components by fitting them in different b-jet multiplicity bins and inclusively in jet multiplicity (to keep the signal contamination low). The components are adjusted by significantly large factors: $t\bar{t}$+$\geq 1c$ is increased by 60%, while $t\bar{t}$+$\geq 1b$ by 30%. Once the normalisation of the components are corrected, a shape correction is also needed. This correction is performed in three steps by weighting the simulated events by values that account for the discrepancy of data to the simulated background in events with two b-jets. This weight is parametrised in terms of different variables to correct for the observed mismodeling of the $t\bar{t}$ process against jet (and large jet) multiplicity, top p_{T} and object angular separation. The modeling of the background after this procedure is significantly improved, as shown in Figure 6b, not only in the agreement of data with the expected background, but also in terms of uncertainty reduction.

A profile likelihood fit is performed in control and signal region simultaneously. The $t\bar{t}t\bar{t}$

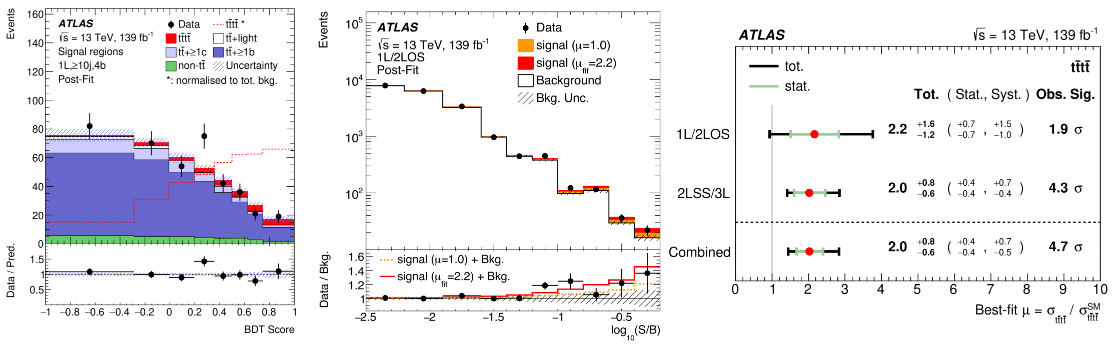
signal strength is fitted together with the constrained normalisation factors of all the $t\bar{t}$ radiation components and the nuisance parameters associated to all the sources of instrumental and modeling systematics. A particular attention is devoted to $t\bar{t}$ modeling uncertainties, consisting of 45 nuisance parameters included in the fit. The control regions contain events with less than 9 (7) and less than 4 b-jets in the 1L (2LOS) channel. The regions at higher b-jet multiplicity with 8 (6) jets are also considered as control regions. In these regions, the shape of H_{T} distribution is fitted. Regions labeled as $3bV$ are not included in the fit, but are considered as validation regions, to validate the background modeling closest to the signal region. The remaining regions are considered as signal regions. In order to maximise the separation of the $t\bar{t}t\bar{t}$ signal to the background, a BDT is trained and its output is used as discriminant variable. The signal strength of the $t\bar{t}t\bar{t}$ process is measured to be 2.2 times the SM value, compatible with the SM hypothesis within one standard deviation and consistent with the result in the SSML channel. An example of the excellent modeling of the data events in the signal region is shown in Figure 7a, where the excess of the data in the most significant bins is also visible.

The results from the two channels are eventually combined. The dominant sources of uncertainties are different in the two channels due to the different sources of the background, therefore the correlation of the systematic uncertainties is not expected to significantly impact in the result. The combined result is dominated by the measurement in the SSML channel and it gives a measured cross-section of 24^{+7}_{-6} fb for the $t\bar{t}t\bar{t}$ process. The gain of the combined result is visible in the observed (expected) significance with respect to the background-only hypothesis, that becomes 4.7 (2.6) standard deviations. As shown in Figure 7b, the excess of the data with respect to background accumulates in the most significant bins for the $t\bar{t}t\bar{t}$ signal and is well in agreement with the fitted SM signal strength. A summary of the results on this measurement is shown in Figure 7c.

4. Conclusions

This document presents the latest results of the ATLAS Collaboration in the measurement of the $t\bar{t}t\bar{t}$ process. The analyses in two different channels are presented as well as the combined result. The results of both analyses are compatible with the SM prediction. Their combination measures a $t\bar{t}t\bar{t}$ cross-section of of 24^{+7}_{-6} fb, compatible with the SM prediction within two standard deviations. The observed (expected) significance of this process over the background is 4.7 (2.6) standard deviations, resulting in the first evidence of this process.

PoS(EPS-HEP2021)489



(a) **(b)** **(c)**

Figure 7: Agreement of the data with expected background in the most significant signal region in the single lepton channel (a). Agreement of the data with the expected background in the analysis bins, ordered by significance of the $t\bar{t}t\bar{t}$ signal (b). Summary of the ATLAS results on the $t\bar{t}t\bar{t}$ production measurement (c) [15].

References

[1] L. Evans and P. Bryant (editors), JINST 3 (2008) S08001.

[2] ATLAS Collaboration, JINST 3 (2008) S08003.

[3] R. Frederix, D. Pagani and M. Zaro, *Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions*, JHEP 02 (2018) 031.

[4] N. Craig, F. D'Eramo, P. Draper et al. , *The hunt for the rest of the Higgs bosons*, JHEP 06 (2015) 137.

[5] N. Craig, J. Hajer, YY. Li, et al., *Heavy Higgs bosons at low* $\tan\beta$*: from the LHC to 100 TeV* , JHEP 01 (2017) 018.

[6] T. Plehn and T. M. P. Tait, *Seeking sgluons*, J. Phys. G: Nucl. Part. Phys. 36 075001.

[7] S. Calvet, B. Fuks, P. Gris et al. , *Searching for sgluons in multitop events at a center-of-mass energy of 8 TeV*, JHEP 04 (2013) 043.

[8] G. R. Farrar and P. Fayet, *Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry*, Phys. Lett. B 76 (1978) 5.

[9] H. P. Nilles, *Supersymmetry, supergravity and particle physics*, Phys. Report 110 (1984) 1-2.

[10] J. A. Aguilar Saavedra, C. Degrande, G. Durieux et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, arXiv:1802.07237 [hep-ph].

[11] QH Cao, SL Chen and Y. Liu, *Probing Higgs width and top quark Yukawa coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ productions*, Phys. Rev. D 95 053004.

[12] QH Cao, SL Chen and Y. Liu et al., *Limiting top quark-Higgs boson interaction and Higgs-boson width from multitop productions*, Phys. Rev. D 99 113003.

[13] ATLAS Collaboration, *evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, EPJC 80 (2020) 1085.

[14] ATLAS Collaboration, *Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector*, ATLAS-CONF-2019-045.

[15] ATLAS Collaboration, *Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, arXiv:2106.11683.