

The measurement of the boosted $t\bar{t}Z$ process

Chaimaa Karam ^{*1}

¹Faculty of Sciences, Mohammed 5 University, Rabat, Morocco

August 25, 2024

This document is prepared as part of the CERN Summer Project 2024 under the supervision of Nedaa-Alexandra Asbah and Knut Zoch.

Abstract

This study delves into the boosted top-quark pair production in association with a Z boson ($t\bar{t}Z$) process in the three-lepton channel. The performance of the next-generation fast simulation tool, AtlFast3 (AF3) [1], is compared with the comprehensive ATLAS full simulation (FS). A significant part of this research is dedicated to verifying the top-tagging efficiency of re-clustered jets (RC Jets) and determining their mass utilizing specific criteria and algorithms.

1 INTRODUCTION

The top quark, the third quark generation with up-charge, was discovered at the Tevatron proton-antiproton collider at Fermilab in 1995 [2, 3]. It is the heaviest particle in the Standard Model (SM) with a mass of approximately 173 GeV [4]. The large Yukawa coupling to the Higgs boson and its mass makes it the crucial player in electroweak (EW) physics. The top quark's unique property of decaying before hadronization allows the study of a bare quark, providing unique insight into fundamental particle interactions [5].

In proton-proton (pp) collisions at the Large Hadron Collider (LHC), top quarks are predominantly produced via the strong interaction, resulting in a top quark-antiquark pair $t\bar{t}$ [5]. At leading order (LO) in QCD, this strong $t\bar{t}$ production can occur via gluon-gluon fusion or quark-antiquark annihilation (see Figure 1), with gluon-gluon fusion being the dominant process at LHC energies [7].

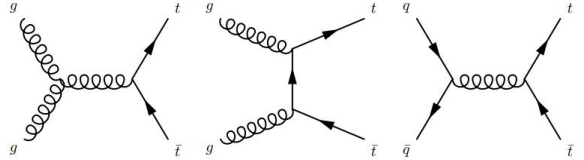


Figure 1: Examples of Feynman diagrams for $t\bar{t}$ production at the LO in QCD include gluon fusion in the s-channel (left) and t-channel (middle), as well as quark-antiquark annihilation (right).

According to the SM, producing a top-quark antiquark pair in association with Z boson $t\bar{t}Z$ is a rare process. The production of a top-quark and anti-top-quark pair in association with a Z boson ($t\bar{t}Z$) is a rare process in the Standard Model. At 13 TeV, the cross-section for this process has been measured to be:

$$\sigma_{t\bar{t}Z} = 860 \pm 60 \text{ fb}$$

combining both statistical and systematic uncertainties. This process is of particular interest be-

^{*}chaimaa_karam@um5.ac.ma

cause it provides direct access to the weak coupling of the top quark to the Z boson. Precise measurement of the inclusive and differential cross-sections of $t\bar{t}Z$ production is essential for testing the SM and exploring potential beyond the Standard Model (BSM) physics [6]. The $t\bar{t}Z$ process can be studied through various final states, including the trilepton (3l) channel and the dilepton (2l) channel. In the 3l channel, events with three leptons are selected with the same-flavor opposite-sign leptons, it provides a good signal-to-background ratio due to the distinct signature of three leptons.

Three main tasks were undertaken in this project. The first task involved comparing AtlFast3, the next generation of fast simulation, with ATLAS full simulation. This comparison aimed to evaluate the accuracy and efficiency of AF3, particularly in reconstructed variables. The second task involved implementing top tagging criteria and rigorously evaluating their effectiveness to ensure they are robust and accurate for identifying top quarks. Top tagging is important for identifying events involving top quarks, especially in the boosted region with high momentum, where the resolved region may fail. The third task focused on calculating the mass of the RC Jet using its sub-jets. This was particularly relevant in the context of the boosted $t\bar{t}Z$ process in the 3l channel. Accurate mass measurements are essential for understanding the jets' properties and distinguishing signals from the background in various analyses. These tasks are integral to improving the precision of simulations and analyses in high-energy physics, contributing to both the validation of the Standard Model and the search for new physics phenomena.

2 Simulation overview

2.1 AltFast3 and Full simulation in ATLAS detector

The ATLAS experiment at the Large Hadron Collider (LHC) relies heavily on accurate simulations of Monte Carlo (MC) to complement the data collected. The full simulation using the Geant4 toolkit (G4) [8] is highly accurate but extremely CPU-

intensive, particularly for simulating the complex geometry of the ATLAS electromagnetic calorimeter. To address these challenges, the ATLAS Collaboration developed AtlFastII [9], a fast simulation tool that combines various methods to simulate different sub-detectors and particles. However, AtlFastII has limitations in accurately modeling jets and their sub-structure.

To overcome these limitations, the ATLAS Collaboration introduced AtlFast3 [1], which maintains the same CPU performance as AtlFastII but offers improved accuracy in reproducing Geant4 simulations. AtlFast3 incorporates advanced parametric simulations like FastCaloSim V2 and FastCaloGAN, enabling the efficient and precise simulation of particle showers, crucial for large-scale resimulation campaigns and future data-taking runs. AtlFast3 is designed to simulate particle showers with such precision that the reconstruction algorithms cannot distinguish any significant differences from the Geant4 simulation. This includes algorithms for reconstructing and identifying electrons, photons, and τ -leptons, as well as for jet reconstruction and clustering. This high level of accuracy ensures that AtlFast3 can be effectively used for detailed physics analyses without compromising on the quality of the simulation.

2.2 Jet re-clustering

Jet re-clustering is a technique that improves the reconstruction of large-radius jets, which is essential for capturing the decay products of boosted heavy particles at the LHC. This method involves first calibrating jets at a smaller radius, typically using radii $R = 0.4$, and then reclustering them at a larger radius, $R = 1.2$. Re-clustering enhances the performance of large-radius jets without requiring additional calibration, making it particularly valuable for precision measurements and searches for new physics [10].

2.3 Event generation and data sample

The production of a top-antitop quark pair along with a Z boson is discussed in [11]. The MadGraph5_aMC@NLO 2.8.1 generator is utilized to

simulate this procedure with Next-to-Leading Order (NLO) precision. Pythia 8.244 manages the parton shower, depicting the progression of quarks and gluons following the initial hard collision. The NNPDF3.0nlo Parton Distribution Function (PDF) set is utilized to depict the momentum distribution of the partons within protons. The $t\bar{t}Z$ process is expected to occur in proton-proton collisions with a likelihood of 876 femtobarns (fb) as indicated by the reference cross-section.

In this work, we define a trilepton preselection, exactly three signal leptons are required, with transverse momenta exceeding 27, 15, and 15 GeV. Among these leptons, the OSSF pair with an invariant mass closest to the Z-boson mass is assumed to originate from the Z-boson decay, and its invariant mass ($m_{\ell\ell}$) must be within 10 GeV of the Z-boson mass. Additionally, all OSSF lepton pairs must have $m_{\text{OSSF}} > 10$ GeV to exclude contributions from low-mass resonances not included in the simulation. At least three jets are necessary, with at least one being b -tagged (with an 85% tagging efficiency). For 0RC, the transverse momentum must be greater than 200 GeV for the leading jet and more than 25 GeV for the others. For 1RC, the leading jet must exceed 50 GeV, with the other three jets over 25 GeV. All trilepton preselections are detailed in table 1.

Variable	Preselection
N_l ($l = e, \mu$)	$= 3$ ≥ 1 OSSF lepton pair with $ m_{\ell\ell} - m_Z < 10$ GeV For all OSSF combinations: $m_{\text{OSSF}} > 10$ GeV
$p_T(l_1, l_2, l_3)$	$> 27, 15, 15$ GeV
N_{jets}	≥ 3
$p_T(\text{jet}_n), n=1..4$	$> 200, 25, 25, 25$ GeV
$N_{b\text{-tagged jets}}$	≥ 1 @ 85%

Table 1: Trilepton Preselection Criteria.

3 RC Jet top tagged

In this part of my work, I focused on evaluating whether RC jets are top-tagged within the Full simulation framework. This involved applying specific criteria to identify top quarks within RC jets.

3.1 Methodology

The top-tagging criteria included several key factors: the number of sub-jets, the mass of the RC jet, its transverse momentum, and the b -tagging status of its sub-jets (see table 2). I develop a C++ code that enforces these criteria by running multiple checks. Initially, it verifies if there are any RC jets available. If the list of RC jets is empty, the function returns false, indicating that the RC jet is not top-tagged. If RC jets are present, the algorithm selects the first RC jet for evaluation. It then checks if the RC jet has fewer than three sub-jets, a mass less than 100 GeV, and a P_t less than 200 GeV. If these conditions are met, the algorithm further examines the sub-jets to ensure that none are b -tagged. If all these criteria are satisfied, the function returns true, confirming that the RC jet is top-tagged. Otherwise, it returns false. This approach guarantees a methodical and precise assessment of RC jets to achieve the main objective of pinpointing top quarks.

Table 2: Top-Tagging Criteria for Rc Jets

Criterion	Condition
N_{Subjets}	≥ 3
Mass	≥ 100 GeV
P_t	≥ 200 GeV
$N_{b\text{-tagged}}$	≥ 1

4 RC Jet mass using sub-jets

The primary purpose of this part of my work is to precisely calculate the mass of a boosted top quark. In such an event, the decay product of the top quark, usually three jets, can be so close together that they are often clustered into a single large jet, known as a re-clustered jet.

4.1 Methodology

I write a code that begins by identifying sub-jets within the RC Jet. These sub-jets are smaller jets within the large RC Jet that represent the top quark's decay products. If the RC Jet contains exactly three sub-jets, these are used directly to calculate the top quark mass. In cases where the RC Jet has fewer than three sub-jets, the code supplements the missing jets by searching the closet jet in the jet collection. It identifies jets that are not part of the initial RC Jet and are close in angular distance (ΔR) to the RC Jet.

Let $\{J_k\}$ be the set of selected additional jets such that:

$$\begin{aligned} \min \Delta R_k &= \Delta R(J_{RC}, J_k) \quad \text{for } J_k \notin \{j_1, j_2, \dots, j_n\} \\ &|\{j_1, j_2, \dots, j_n\} \cup \{J_k\}| = 3 \end{aligned} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

For this case we have ($n < 3$), where $\{J_{RC}\}$ is the RC Jet and $\{j_1, j_2, \dots, j_n\}$ is the set of sub-jets within J_{RC} . This equations (1) and (2) are to ensure that three jets are always available for the mass calculation.

If the RC Jet has over three sub-jets, we refine the selection by categorizing the sub-jets into b-tagged jets, originating from bottom quark, and light jets. The priority is given to the b-tagged jets, its sort by their transverse momentum, and the top three jets are selected for mass reconstruction.

Once three jets are chosen, their momenta are combined to calculate the invariant mass of the RC jet, used as reconstructed mass, is given by

$$m_{RCjet} = \sqrt{(E_1 + E_2 + E_3)^2 - (\vec{p}_1 + \vec{p}_2 + \vec{p}_3)^2} \quad (3)$$

where E_i is the energy and \vec{p}_i is the three-momentum of the i -th jet.

5 Results

The first part of my work aimed to compare two simulations, AtlFast3, and FS, in the boosted ttZ process within a region involving three leptons. This comparison focused on two categories: 0RC (no re-clustered jet) and 1RC (one re-clustered jet). Several

key variables were analyzed, including the transverse momentum (p_T) of the leading jet and the next three jets, as well as the p_T of the leading, second, and third leptons. The study also looked at contributions from fake leptons and charge flips, missing transverse energy (MET), the invariant mass of the lepton pair ($m_{\ell\ell}$), the transverse mass of the W boson (m_T^W), and the number of b-tagged jets, total jets, and electrons in the event. In the 1RC category, additional features of the leading re-clustered jet were examined, including its mass, the number of b-tagged sub-jets and total sub-jets, and the p_T of both the leading re-clustered jet and its first, second, and third sub-jets. The results from Figure 2-25 indicate that AF3 provides reasonable results, comparable to those of FullSim. Both simulations produce almost similar event counts for the given bins, and the ratios are very close to 1, indicating that AF3 provides a similar response to FS. However, AtlFast3 requires calibration to ensure its outputs are accurate and reliable, closely aligning with those of FS. This calibration is crucial for maintaining consistency and reliability in the analysis of the boosted ttZ process. Second, I focused on evaluating if the RC jet is top-tagged. The result shown in Figure 26 indicates that more events in 1RC are top-tagged. In the third part, figure 27 shows that the RC jet mass distribution is centered around 207.4 GeV, exhibiting a wide range of values. Most jets have a mass between 150-200 GeV, with fewer jets exceeding 300 GeV in 1RC. From Figures 28 -30, most events contain fewer than 3 sub-jets, with fewer events having exactly 3 sub-jets. It is rare to observe events with more than 3 sub-jets.

6 Conclusion

This project compared AF3 and FS simulations in the boosted ttZ process with three leptons, focusing on 0RC and 1RC categories. The results show that AF3, with proper calibration, provides results comparable to FS. Additionally, more events in the 1RC category are top-tagged, and the RC jet mass distribution is centered around 207.4 GeV. Most events contain fewer than 3 sub-jets.

Acknowledgment

I am grateful to CERN for giving me the chance to participate as a summer student in 2024 and to attend classes taught by brilliant professors. I appreciate the guidance and support provided by my supervisors Nedaa-Alexandra Asbah and Knut Zoch throughout this project. I'm very grateful to my parents for their support, as their faith in me has always given me inner strength. Finally, I'd like to thank my professors at Université Mohammed 5 for their advice and support throughout my master's studies.

References

- [1] ATLAS Collaboration, “AtlFast3: the next generation of fast simulation in ATLAS,” *Computing and Software for Big Science*, vol. 6, no. 1, p. 7, 2022.
- [2] D0 Collaboration, “Observation of the top quark,” *Physical Review Letters*, vol. 74, pp. 2632–2637, 1995. arXiv:hep-ex/9503003.
- [3] CDF Collaboration, “Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab,” *Physical Review Letters*, vol. 74, pp. 2626–2631, 1995. arXiv:hep-ex/9503002.
- [4] ATLAS Collaboration *et al.*, “Measurement of the top quark mass in the $t\bar{t} \rightarrow$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data,” *arXiv preprint arXiv:1606.02179*, 2016.
- [5] F. Déliot and P. Van Mulders, “Comptes Rendus Physique,” *Comptes Rendus Physique*, 2020.
- [6] O. B. Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou, and C. Zhang, “Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD,” *Journal of High Energy Physics*, vol. 2016, no. 5, pp. 1–39, 2016.
- [7] ATLAS Collaboration, “Climbing to the Top of the ATLAS 13 TeV data,” *Journal/Conference Name*, vol. 2024.
- [8] GEANT Collaboration, S. Agostinelli, et al., *GEANT4—a simulation toolkit*, Nucl. Instrum. Meth. A, vol. 506, no. 1-2, pp. 250-303, 2003.
- [9] ATLAS Collaboration, “Fast simulation for ATLAS: Atlfast-II and ISF,” *Journal of Physics: Conference Series*, vol. 396, no. 2, pp. 022031, 2012.
- [10] Nachman, Benjamin and Nef, Pascal and Schwartzman, Ariel and Swiatlowski, Maximilian and Wanotayaroj, Chaowaroj, “Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC”, *Journal of High Energy Physics*, vol. 2015, no. 2, pp. 1–18, 2015.
- [11] ATLAS Collaboration, “Inclusive and differential cross-section measurements of $t\bar{t}Z$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, including EFT and spin-correlation interpretations,” *Journal of High Energy Physics*, 2024.

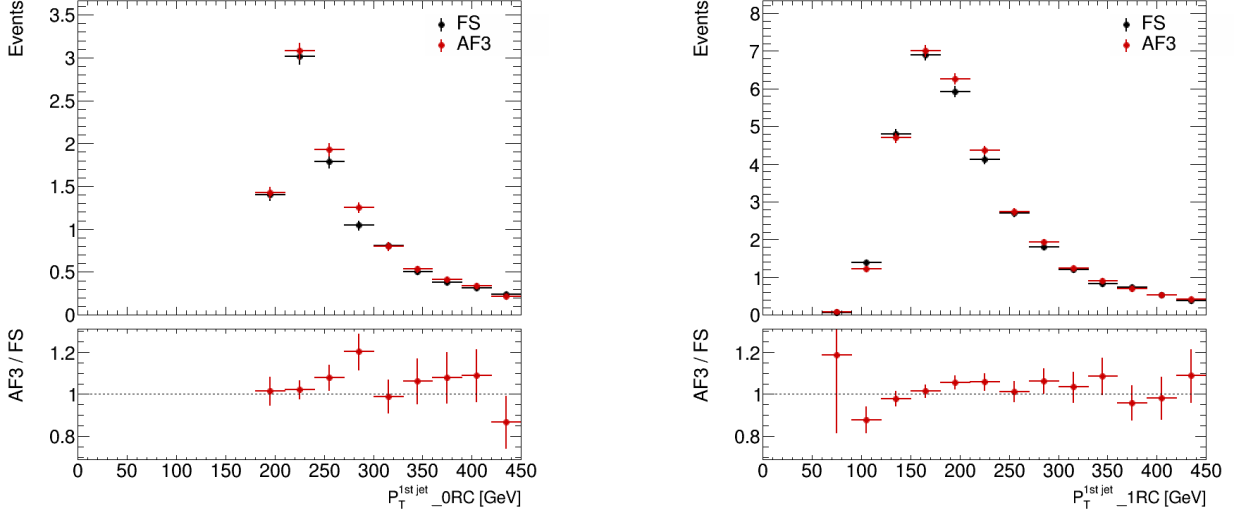


Figure 2: p_T distributions for the leading jet in 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations

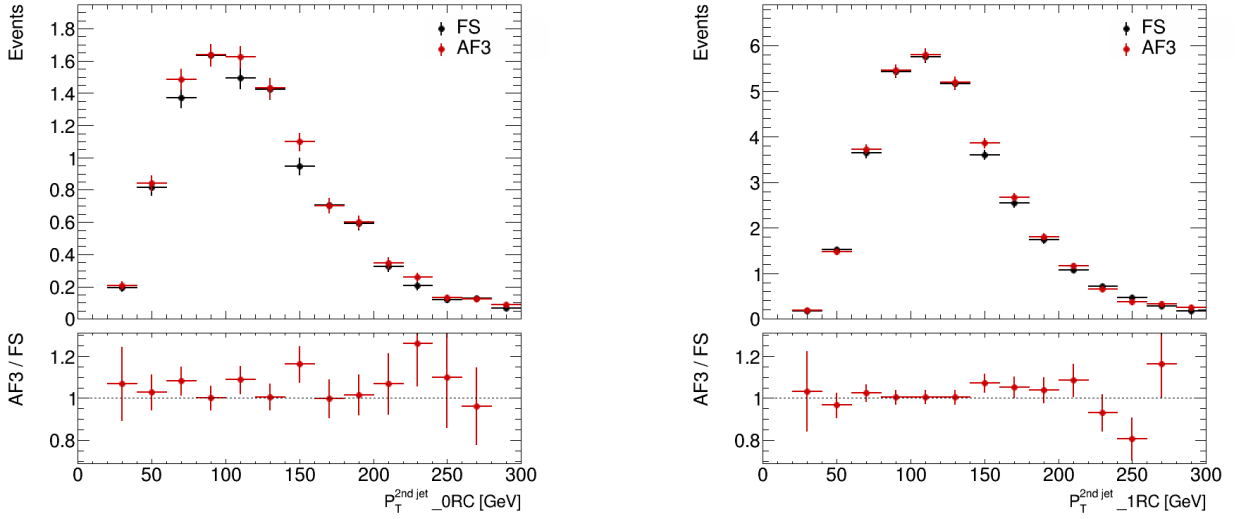


Figure 3: p_T distributions for the second jet in 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

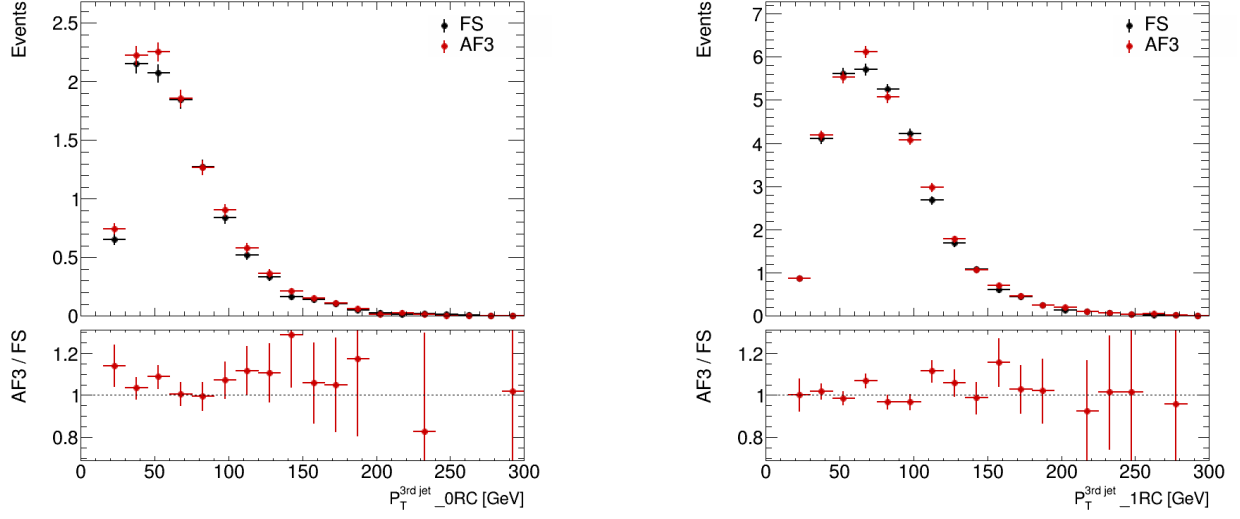


Figure 4: p_T distributions for the third jet in 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations

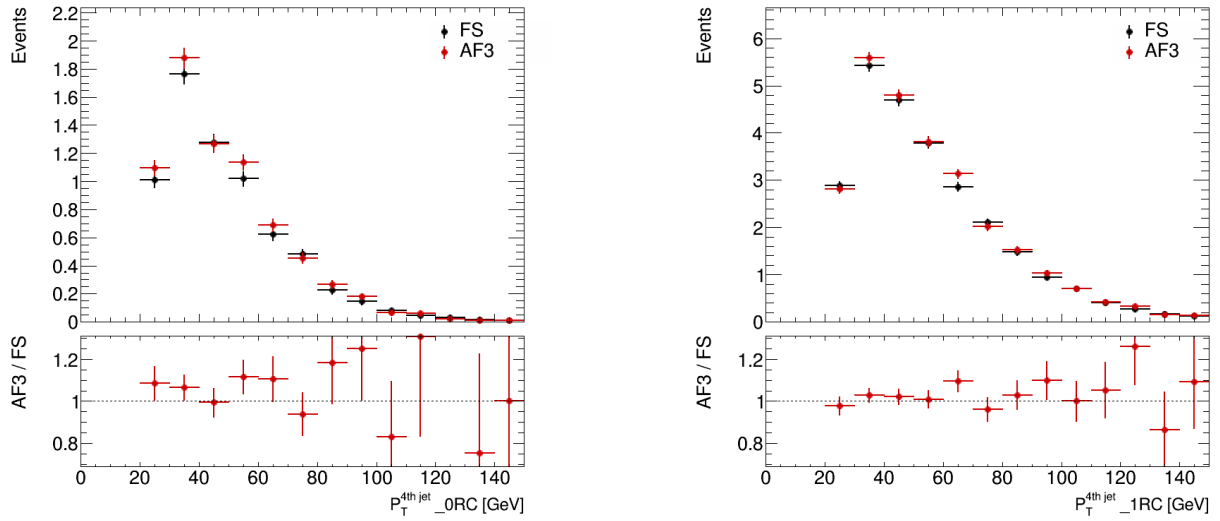


Figure 5: p_T distributions for the fourth jet in 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

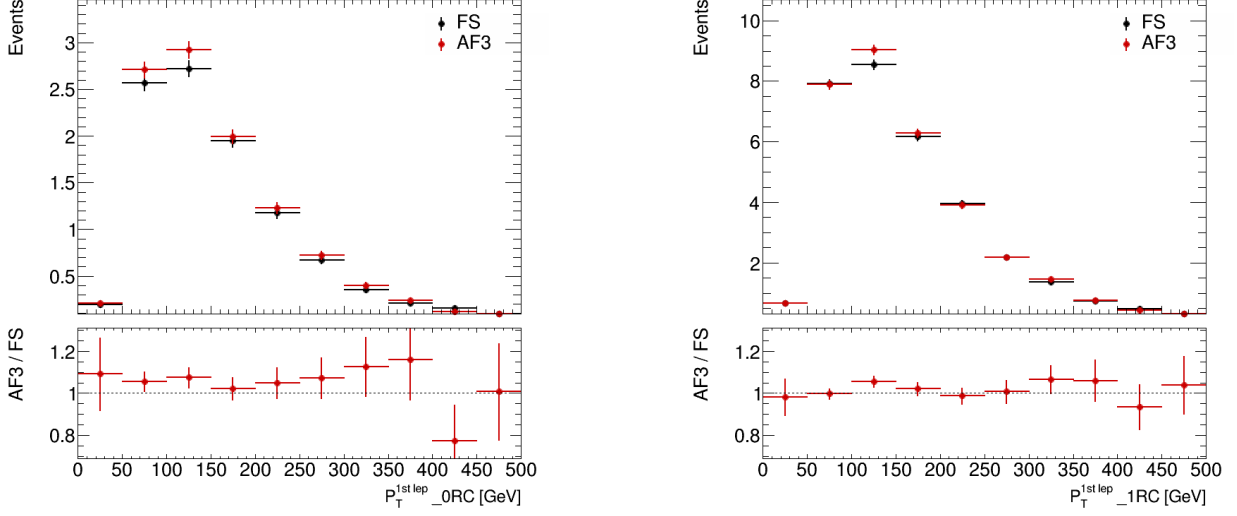


Figure 6: p_T distributions for the leading lepton in 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

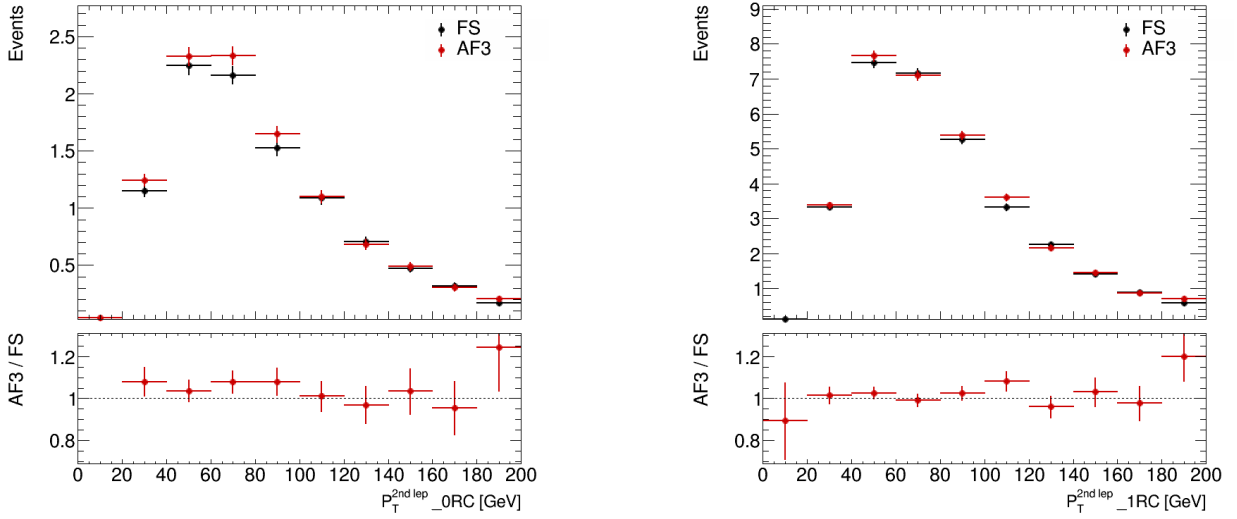


Figure 7: p_T distributions for the second lepton in 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

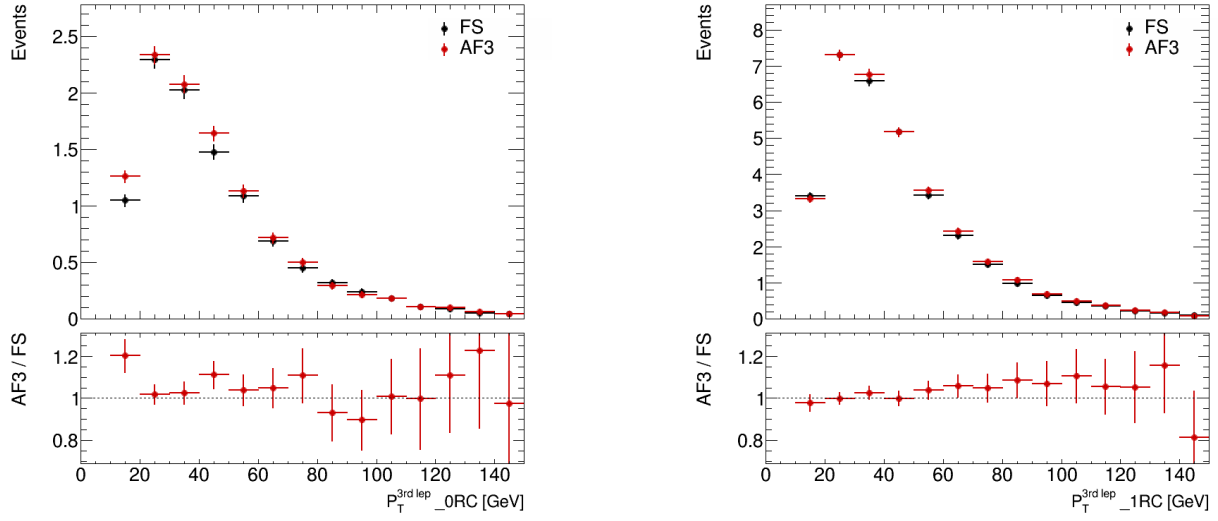


Figure 8: p_T distributions for the third lepton in 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

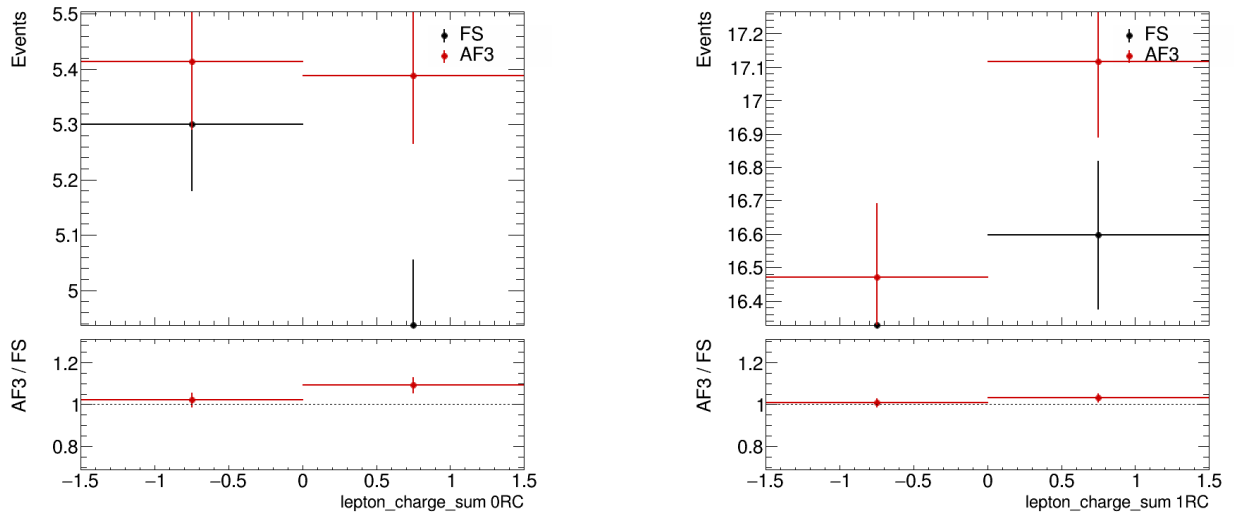


Figure 9: The sum of lepton charges was calculated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

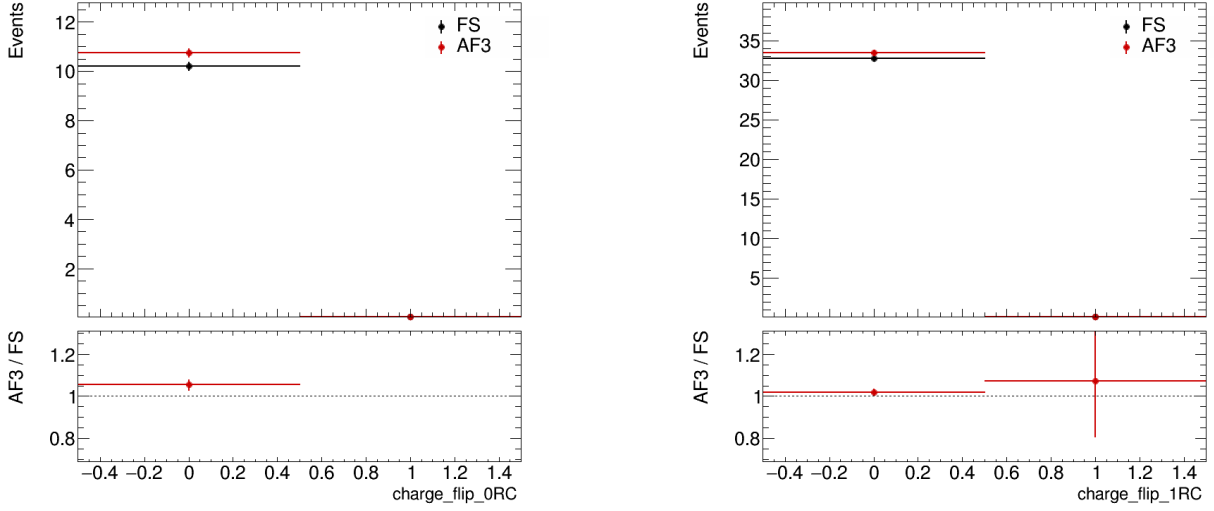


Figure 10: The charge flip in leptons was analyzed for both 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

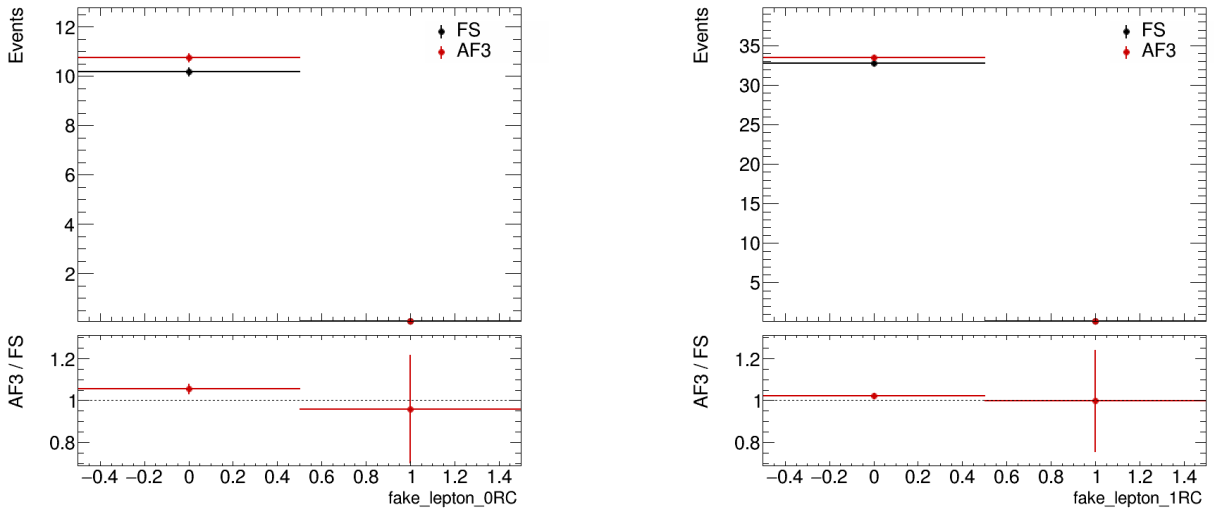


Figure 11: The fake lepton was analyzed for both 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

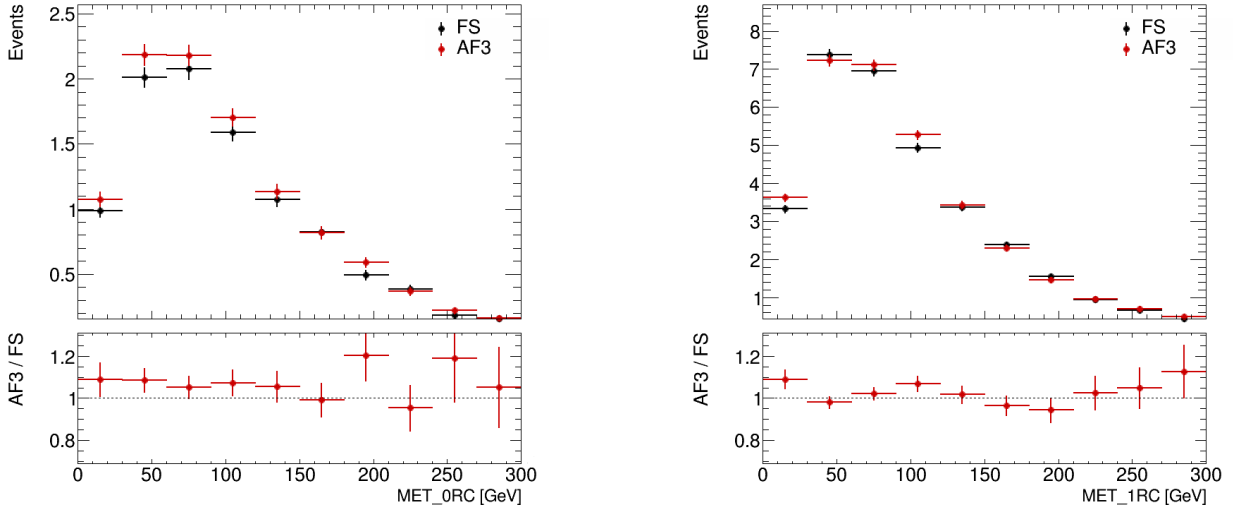


Figure 12: The missing transverse energy was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

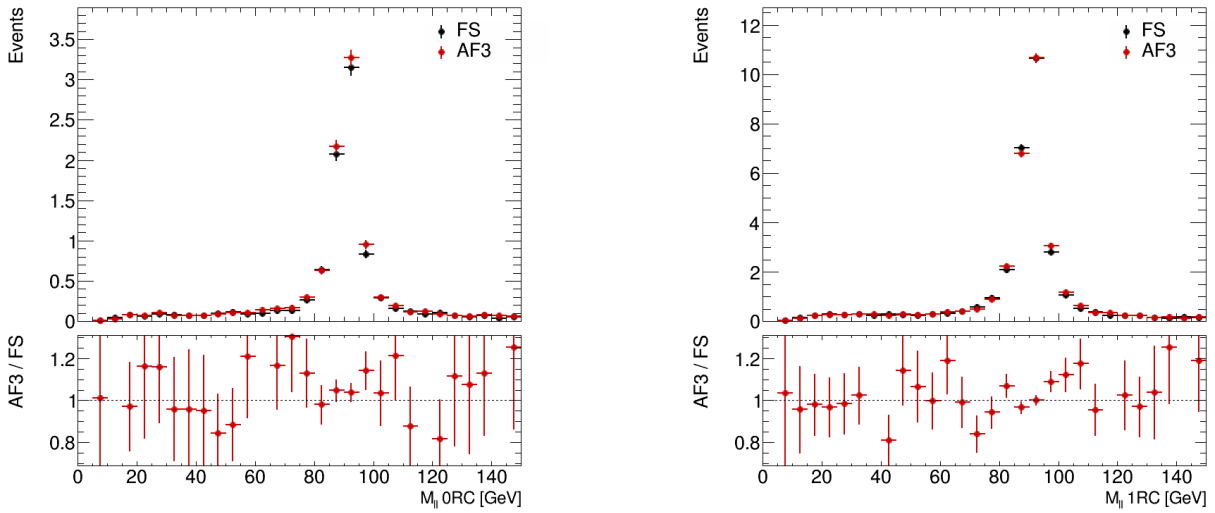


Figure 13: the invariant mass of a lepton pair was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

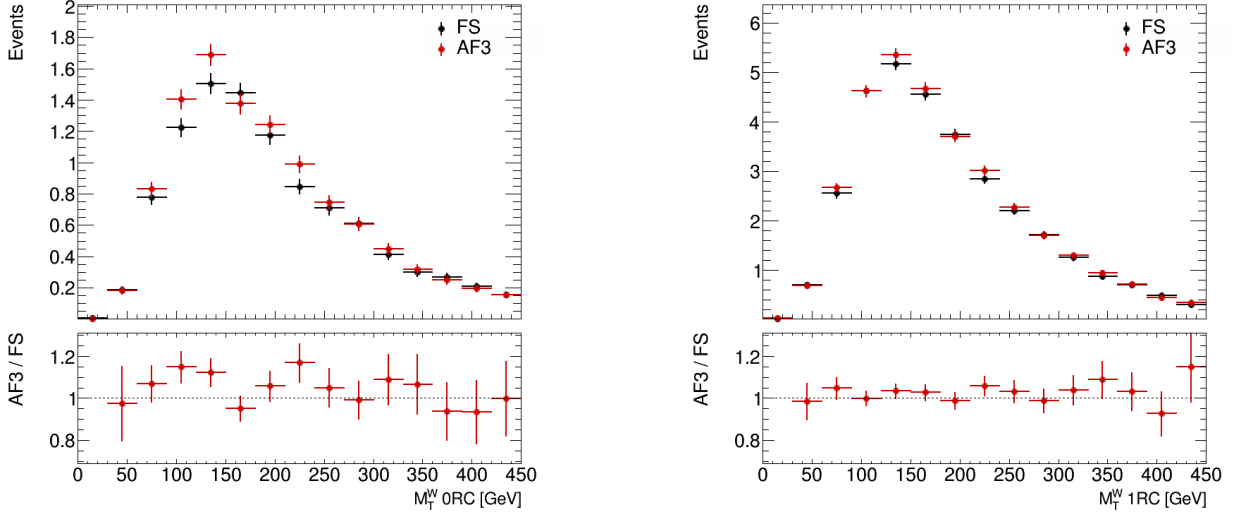


Figure 14: the transverse mass of the W boson was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

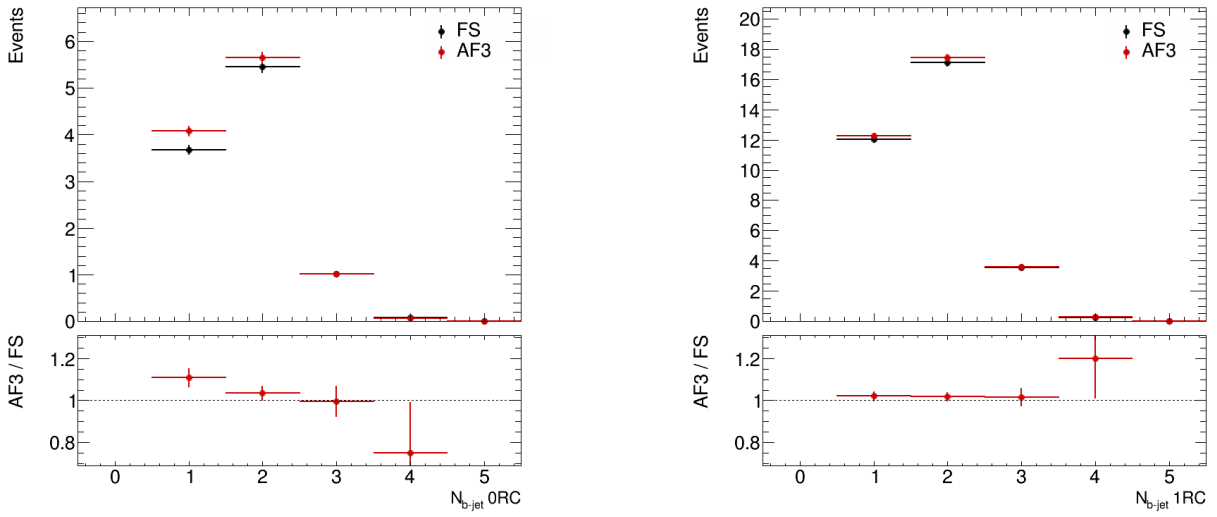


Figure 15: The number of b jets was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

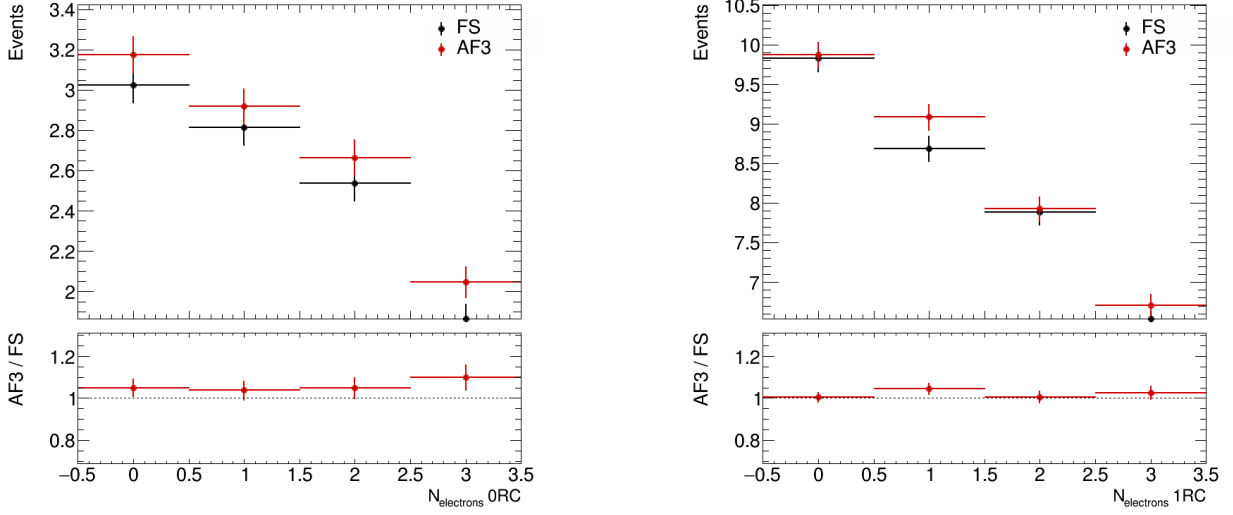


Figure 16: The number of electrons was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtIFast3 and FS detector simulations.

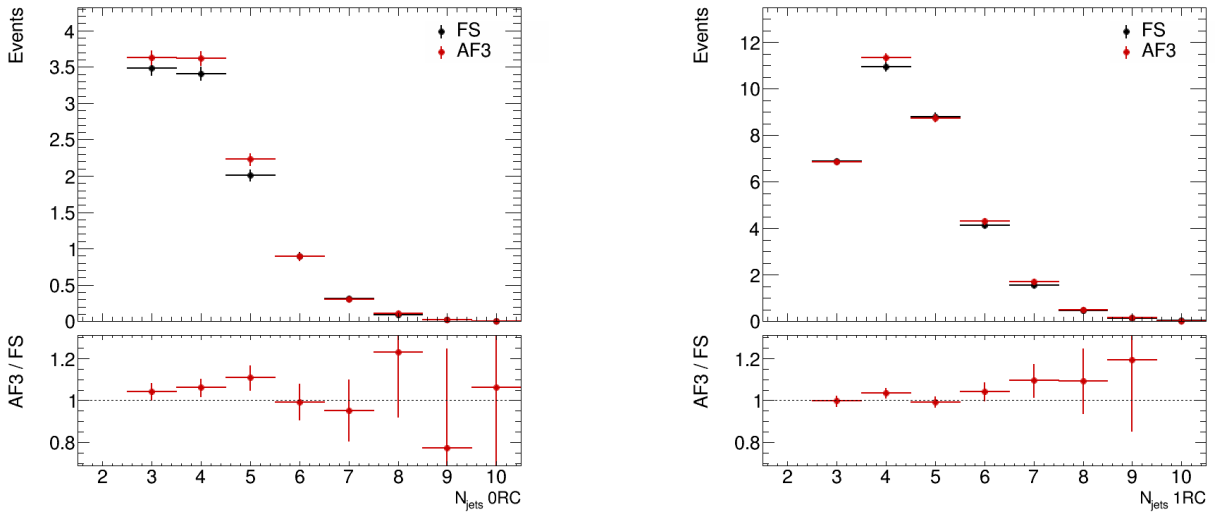


Figure 17: The number jets was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtIFast3 and FS detector simulations.

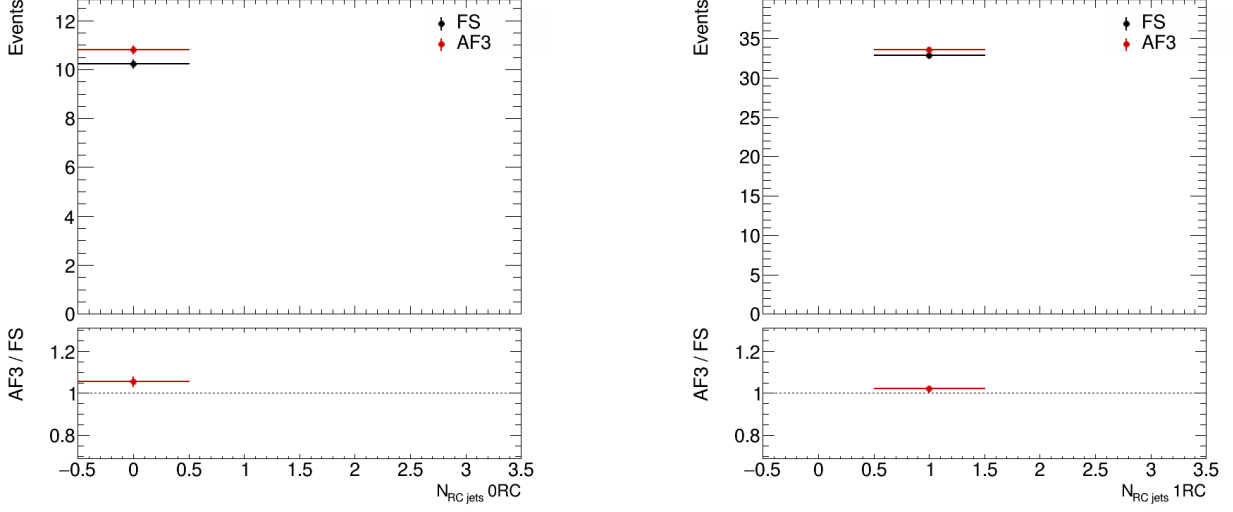


Figure 18: The number RC jets was evaluated at both 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

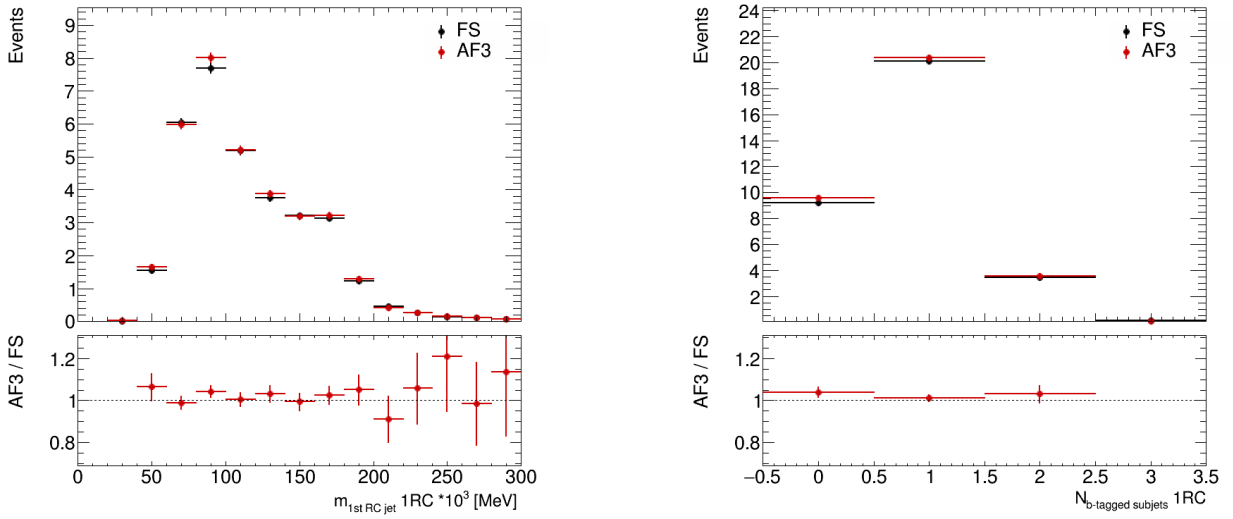


Figure 19: The mass of the leading RC jets (left) and number of b-tagged subjects (right) in 1RC using ttZ samples generated with AtI Fast3 and FS detector simulations.

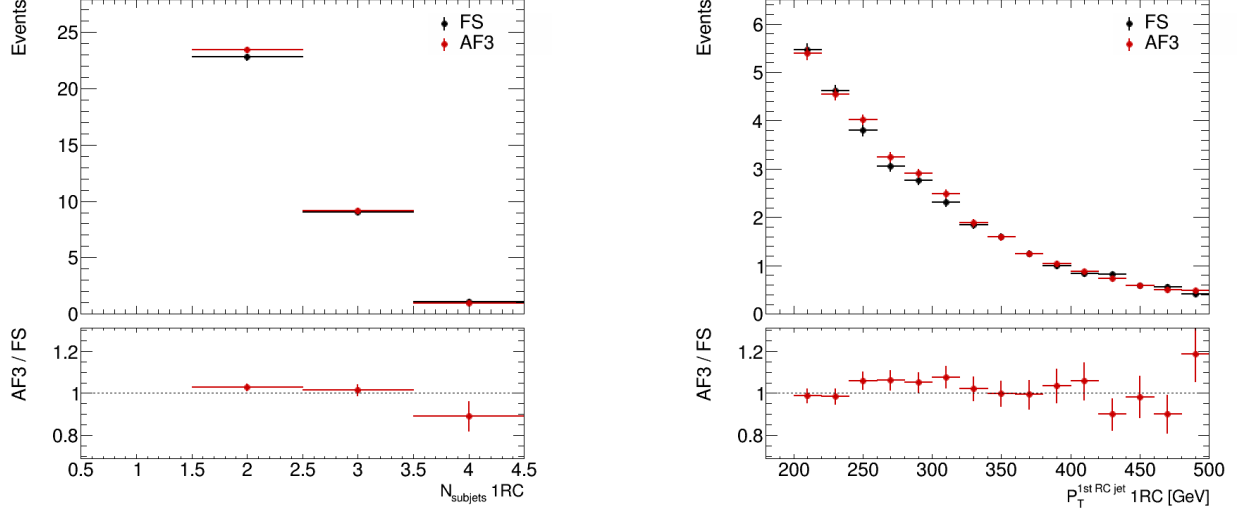


Figure 20: The number of subject in 1RC (left) and P_T distribution of the leading RC jet (right) using ttZ samples generated with AtFast3 and FS detector simulations.

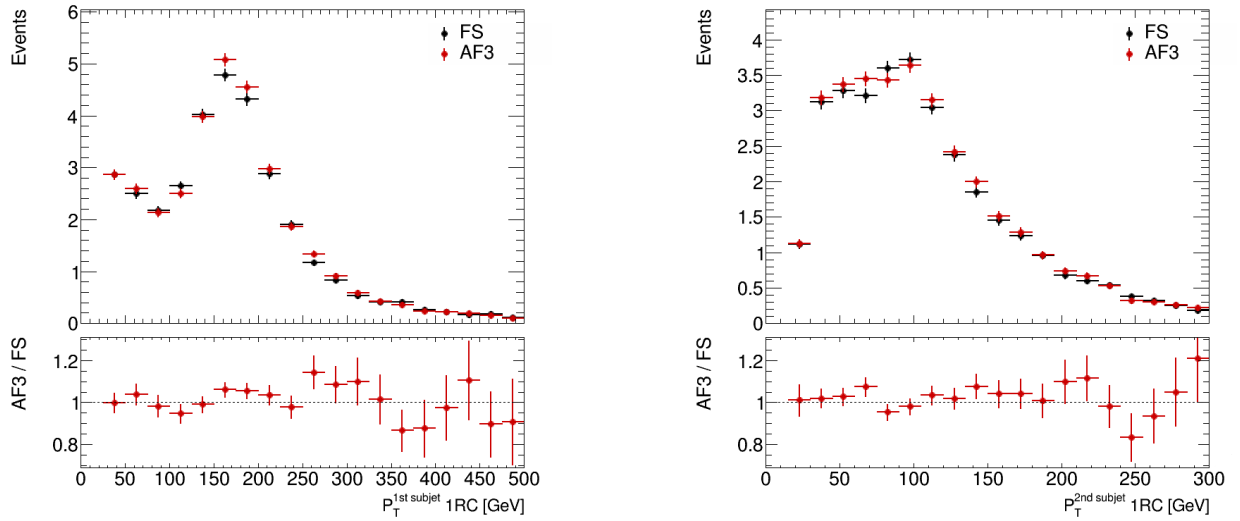


Figure 21: p_T distributions for the leading subject (left) and the second subject (right) in RC jet using ttZ samples generated with AtFast3 and FS detector simulations.

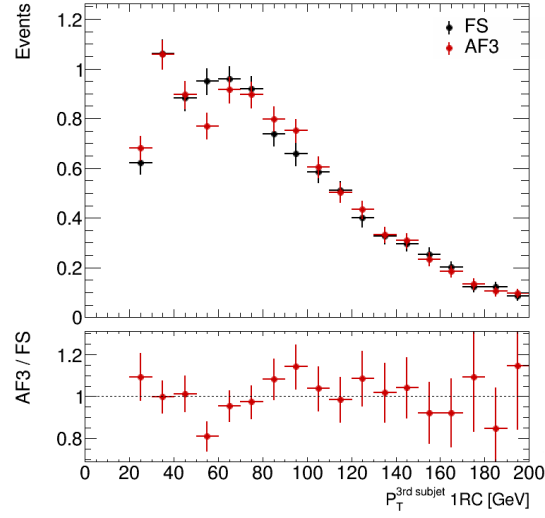


Figure 22: p_T distributions for the third sub-jet in RC jet using ttZ samples generated with AtI Fast3 and FS detector simulations

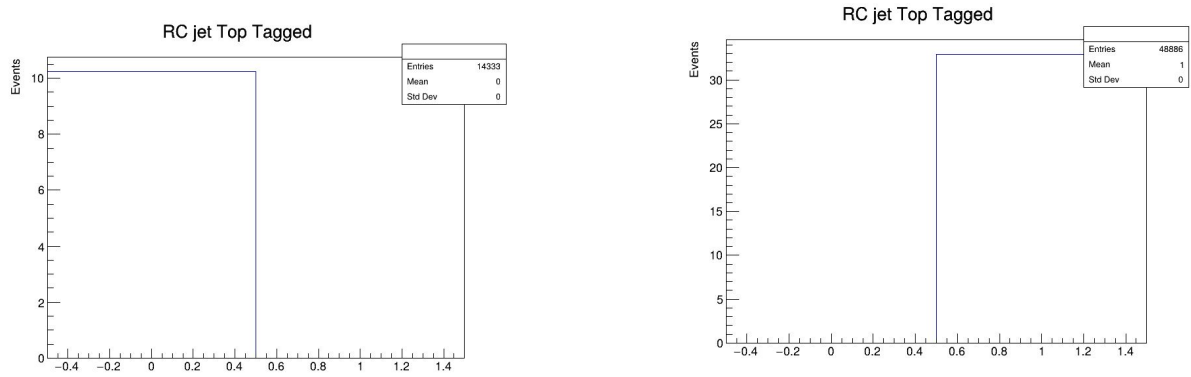


Figure 23: Rc jet Top tagged was evaluated in 0RC (left) and 1RC (right) using ttZ samples generated with AtI Fast3 and FS detector simulations.

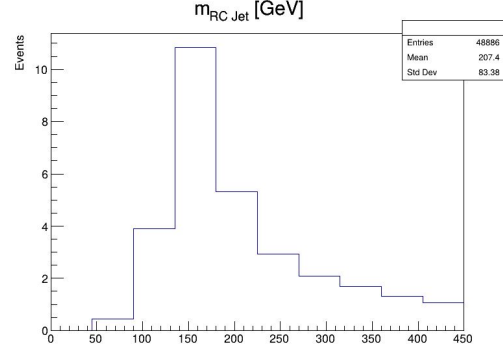
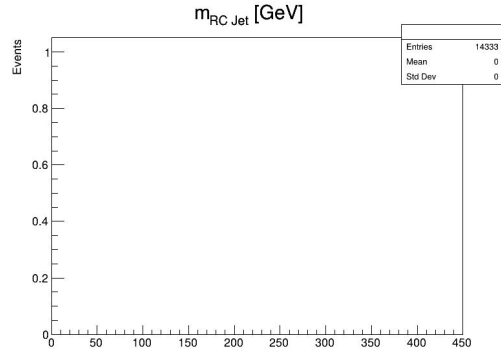


Figure 24: the mass of RC Jet using sub-jets was evaluated in 0RC (left) and 1RC (right) using ttZ samples generated with AtlFast3 and FS detector simulations.

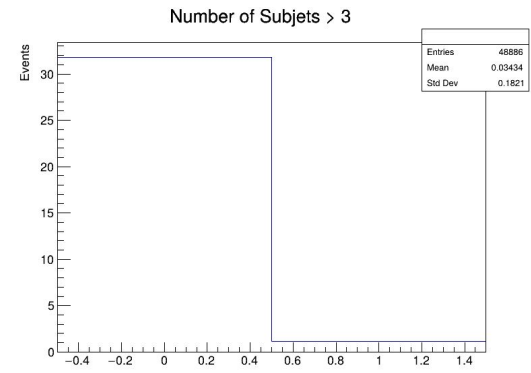
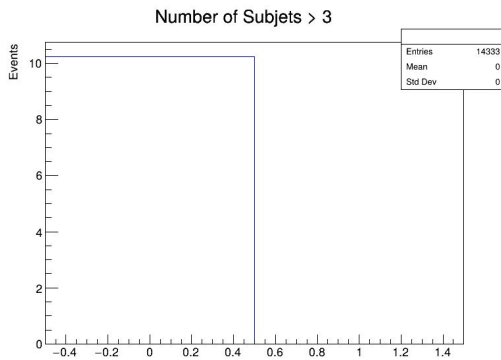
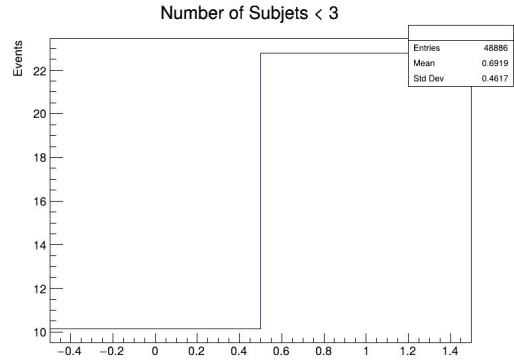
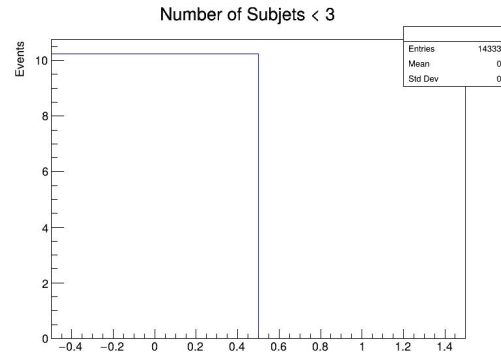


Figure 25: The number of sub-jet fewer than 3 and more than 3 in 0RC and 1RC using ttZ samples generated with AtlFast3 and FS detector simulations.

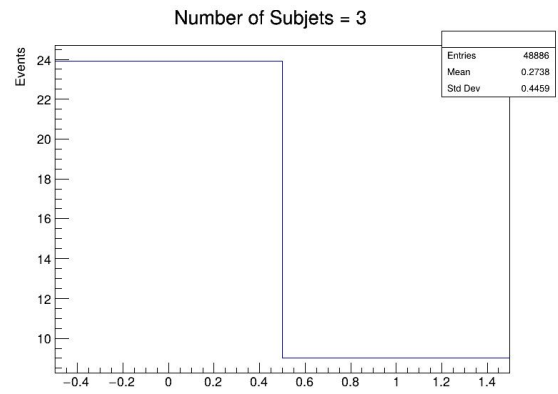
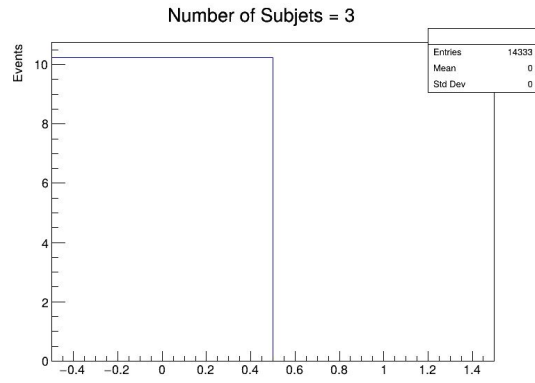


Figure 26: he number of sub-jet exactly 3 in 0RC and 1RC using ttZ samples generated with AtI Fast3 and FS detector simulations.