Study of the difference between substructure variables of large-R jets using AFII and FS simulations

CERN 2020 Summer online learning initiative

Luis Coelho
CERN online learning initative Student, LIP-Coimbra

Nedaa-Alexandra Asbah and Jelena Jovicevic
CERN

(September 25, 2020)

I. INTRODUCTION

With the discovery of a 125 GeV Higgs boson and the rapid progress made at the Large Hadron Collider (LHC), it is essential to precisely measure the properties of the Higgs boson as any potential disagreement could hint at fascinating new phenomena. The Yukawa coupling of the Higgs boson to the top quark is a key parameter of the Standard Model (SM). It can be determined from the ratio of the top quark mass and Higgs field vacuum expectation value, from the cross section of $gg \to H$ production through a top quark loop, or from the cross section of the process $gg/qq \to t\bar{t}H$ [1, 2], which is a tree-level process at lowest order in perturbation theory. Comparison of these measurements has the potential to identify and disambiguate new physics effects that can modify the $t\bar{t}H$ production cross section relative to the SM expectation.

Discrepancies in the high-momentum Higgs boson production in the presence of new physics are predicted to be as high as 50% [3] with respect to the SM expectations. As in any high-energy physics analysis, a precise simulation of the detector response, with an accurate description of the interactions between the particles with the detector components, is essential. However, due to the large number of events demanded and large LHC luminosity, a full detector simulation can be challenging and extremely time-consuming. In some cases, the use of a simplified and fast approach may be a satisfactory alternative. In this work, a study was performed to compare the fast detector simulation Atlfast-II (AFII) [4] with the ATLAS full simulation (FS), and verify the modelling of substructure variables of large-R jets using $t\bar{t}H$ and $t\bar{t}$ Monte Carlo (MC) samples.

II. ATLFAST-II AND FULL DETECTOR SIMULATION

The full ATLAS detector simulation strategy [5] uses a highly detailed detector description and is based on the Geant4 toolkit [6]. Such a detailed simulation requires a large CPU time for each event, which is mostly spent inside the calorimeter systems. The ATLAS collaboration has developed the Atlfast-II to achieve a fast and precise detector simulation within the computing limits of the collaboration. The Atlfast-II applies the FastCaloSim package [7, 8], which reduces the calorimeter simulation time to a few seconds per event, using a parametrized simulation of the particle energy response and the energy distribution in the ATLAS calorimeter. For the inner detector and the muon system, the Atlfast-II uses the full Geant4 simulation.

III. SAMPLES

In this work, $t\bar{t}H$ and $t\bar{t}$ MC samples were produced using both the full ATLAS detector simulation and Atlfast-II.

The production of $t\bar{t}H$ events is modelled in the five flavour scheme using the PowhegBox [9–12] generator at NLO with the `NNPDF3.0nlo` [13] PDF set. The h_{damp} parameter, which regulates the high-momentum emission against which the $t\bar{t}$ system recoils, is set to $3/4 \cdot (m_t + m_{\bar{t}} + m_H) = 352.5$ GeV, where m_t, $m_{\bar{t}}$ and

m_H are the masses for the top-quark, anti-top quark and Higgs boson, respectively. The functional form of the renormalisation and factorisation scale is set to $\sqrt[3]{m_{\mathrm{T}}(t) \cdot m_{\mathrm{T}}(\bar{t}) \cdot m_{\mathrm{T}}(H)}$, where m_{T} is the transverse mass of the generated particle, defined as the square root of the quadratic sum between the particle mass and transverse momentum. The events are showered with Pythia8.230 [14].

The production of $t\bar{t}$ events is modelled using the PowhegBox v2 [15] generator at NLO in QCD in the five flavour scheme, and using Pythia8.230. The h_{damp} parameter is set to $3/4 \cdot (m_t + m_{\bar{t}}) = 1.5\ m_{\mathrm{top}}$ [16], and the functional form of the renormalisation and factorisation scale is set to the default scale $\sqrt{m_{\mathrm{top}}^2 + p_{\mathrm{T}}^2}$.

IV. LARGE-RADIUS JETS AND JET SUBSTRUCTURE

When a particle that has a high transverse momentum (boosted regime) and decays hadronically into several daughter particles, the angular separation between the decay products will be small. A single larger-radius jet referred to as large-R jet can be build containing all of the decay products. The large-R jet is used to mimic the kinematics of the original particle. This approach is more efficient than reconstructing all of the products as individual jets. Several observables can be constructed in order to characterise the internal structure of the large-R jet and its components. These variables are named substructure variables [17] and are usually used to identify the original particle that produced the large-R jet.

V. RESULTS

The first set of studied variables contains basic kinematic variables such as the transverse momentum (p_{T}), the pseudorapidity (η), the azimuthal angle (ϕ), the energy (e) and the invariant mass (m) of the leading and subleading large-R jets, as well as the angular distance (ΔR) between the two jets. Figure 1 to 11 show the AFII and FS distributions of these variables for the leading and subleading large-R jets. Small differences can be seen, especially for the low values of p_{T} and energy of the leading and subleading jets using the $t\bar{t}H$ samples, as well as for the tails of these distributions. Overall, the kinematic distributions studied present a mean error smaller than approximately 25% between AFII and FS.

In addition to the kinematic variables, several jet substructure variables were studied and are shown in Figure 12 to 38. These variables include the n-Subjettiness (τ_1^{wta}, τ_2^{wta}, τ_3^{wta}) [18, 19], ratios of n-Subjettiness (τ_{32}^{wta}, τ_{21}^{wta}), energy correlation functions (ECF_1, ECF_2, ECF_3) [20], energy correlations ratios ($D2$, $C2$, $E3$) [21], the K_t distance between the subjets during merging ($Split_{12}$, $Split_{23}$) [22] and the minimum pair invariant mass (Q_{w}) of the three subjets identified at the splitting scale d_{23} [23]. Considerable differences in the shapes of the distributions of these variables can be seen, especially for the n-Subjettiness and its ratios. Overall the mismodelling appears to be smaller for the samples of $t\bar{t}$.

VI. CONCLUSION

In this work, a comparison between the modelling of $t\bar{t}$ and $t\bar{t}H$ samples produced using the fast detector simulation Atlfast-II and the ATLAS full simulation was performed. Several large-R jets kinematic and substructure variables were analysed, and the differences presented.

[1] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-076 (CERN, Geneva, 2017).

[2] *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-077 (CERN, Geneva, 2017).

[3] Christophe Grojean, Ennio Salvioni, Matthias Schlaffer, and Andreas Weiler, *"Very boosted Higgs in gluon fusion,"* JHEP **05**, 022 (2014), arXiv:1312.3317 [hep-ph] .

[4] W Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, Tech. Rep. ATL-SOFT-PROC-2012-065 (CERN, Geneva, 2012).

[5] G. Aad *et al.* (ATLAS), *"The ATLAS Simulation Infrastructure,"* Eur. Phys. J. C **70**, 823–874 (2010), arXiv:1005.4568 [physics.ins-det] .

[6] S. Agostinelli *et al.* (GEANT4), *"GEANT4–a simulation toolkit,"* Nucl. Instrum. Meth. A **506**, 250–303 (2003).

[7] ATLAS collaboration, *"The ATLAS calorimeter simulation FastCaloSim,"* J. Phys. Conf. Ser. **331**, 032053 (2011).

[8] Coll ATLAS, M Beckingham, M Duehrssen, E Schmidt, M Shapiro, M Venturi, J Virzi, I Vivarelli, M Werner, S Yamamoto, and T Yamanaka, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013 (CERN, Geneva, 2010).

[9] Paolo Nason, *"A New method for combining NLO QCD with shower Monte Carlo algorithms,"* JHEP **11**, 040 (2004), arXiv:hep-ph/0409146 .

[10] Stefano Frixione, Paolo Nason, and Carlo Oleari, *"Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,"* JHEP **11**, 070 (2007), arXiv:0709.2092 [hep-ph] .

[11] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, *"A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX,"* JHEP **06**, 043 (2010), arXiv:1002.2581 [hep-ph] .

[12] Heribertus B. Hartanto, Barbara Jager, Laura Reina, and Doreen Wackeroth, *"Higgs boson production in association with top quarks in the POWHEG BOX,"* Phys. Rev. D **91**, 094003 (2015), arXiv:1501.04498 [hep-ph] .

[13] Richard D. Ball *et al.* (NNPDF), *"Parton distributions for the LHC Run II,"* JHEP **04**, 040 (2015), arXiv:1410.8849 [hep-ph] .

[14] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands, *"An introduction to PYTHIA 8.2,"* Comput. Phys. Commun. **191**, 159–177 (2015), arXiv:1410.3012 [hep-ph] .

[15] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi, *"A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,"* JHEP **09**, 126 (2007), arXiv:0707.3088 [hep-ph] .

[16] *Studies on top-quark Monte Carlo modelling for Top2016*, Tech. Rep. ATL-PHYS-PUB-2016-020 (CERN, Geneva, 2016).

[17] "Jet attributes for run 2 twiki page," `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/Run2JetMoments`.

[18] Jesse Thaler and Ken Van Tilburg, *"Identifying Boosted Objects with N-subjettiness,"* JHEP **03**, 015 (2011), arXiv:1011.2268 [hep-ph] .

[19] Daniele Bertolini, Tucker Chan, and Jesse Thaler, *"Jet Observables Without Jet Algorithms,"* JHEP **04**, 013 (2014), arXiv:1310.7584 [hep-ph] .

[20] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler, *"Energy Correlation Functions for Jet Substructure,"* JHEP **06**, 108 (2013), arXiv:1305.0007 [hep-ph] .

[21] Andrew J. Larkoski, Ian Moult, and Duff Neill, *"Power Counting to Better Jet Observables,"* JHEP **12**, 009 (2014), arXiv:1409.6298 [hep-ph] .

[22] Simone Marzani, Gregory Soyez, and Michael Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, Vol. 958 (Springer, 2019) arXiv:1901.10342 [hep-ph] .

[23] Jessie Shelton, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales* (2013) pp. 303–340, arXiv:1302.0260 [hep-ph] .

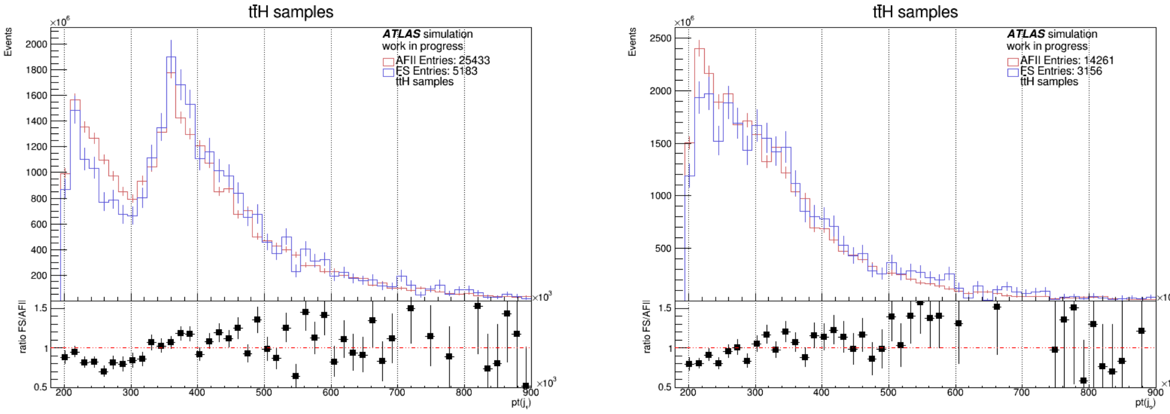

FIG. 1: p_T distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

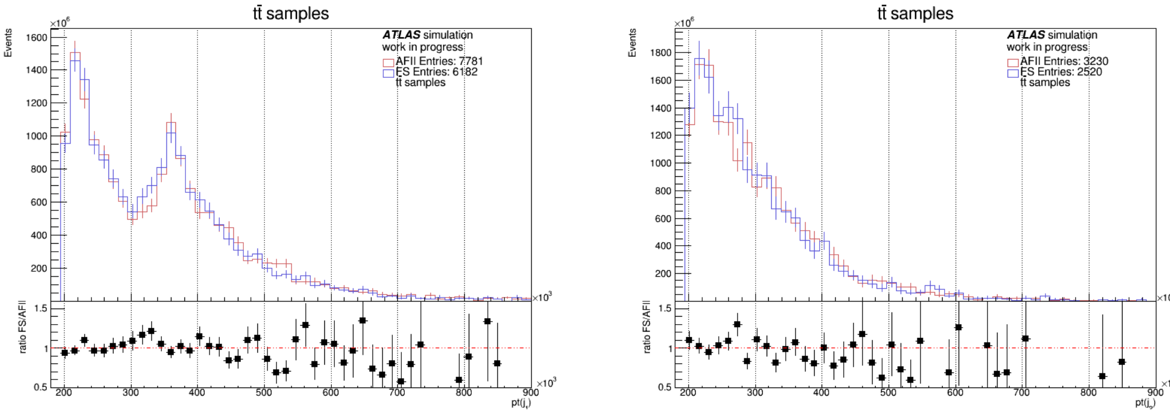

FIG. 2: p_T distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

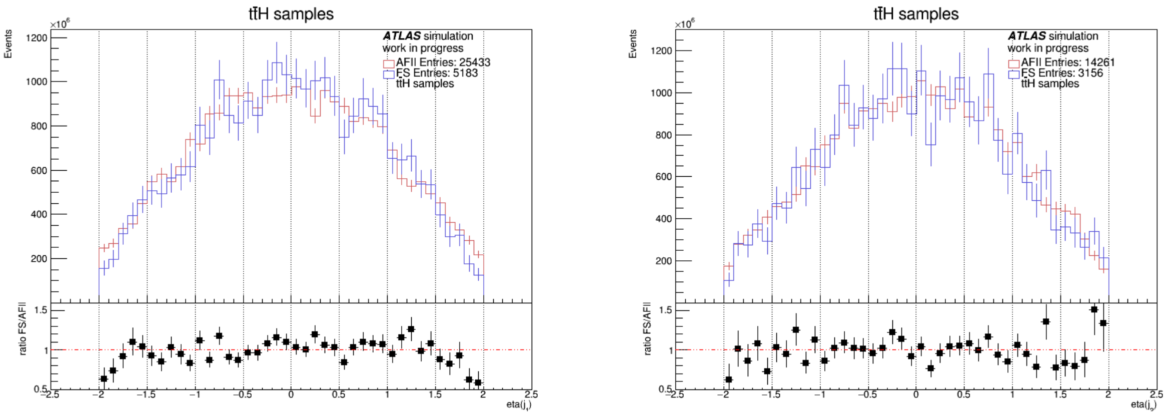

FIG. 3: η distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

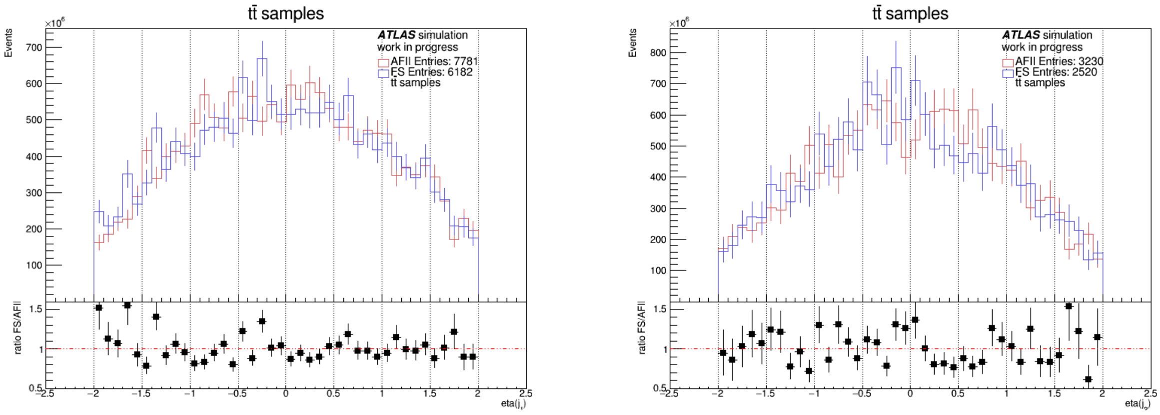

FIG. 4: η distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

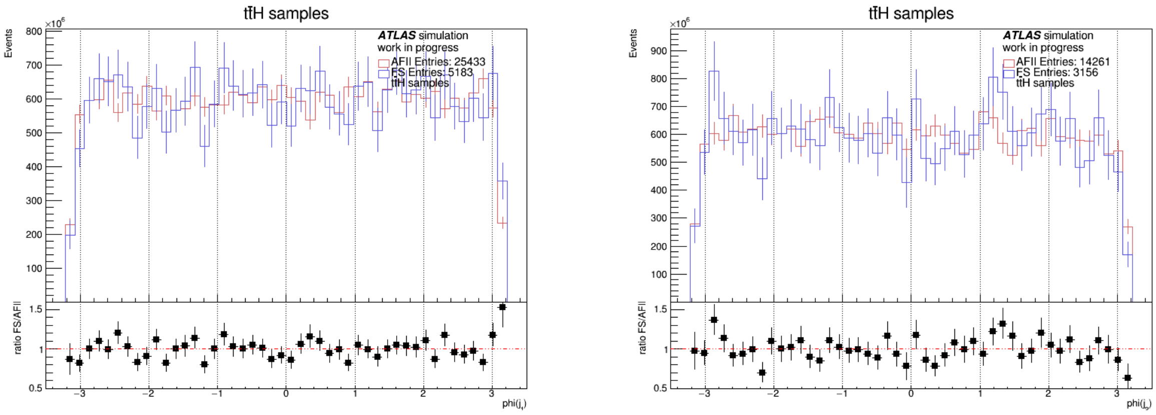

FIG. 5: ϕ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

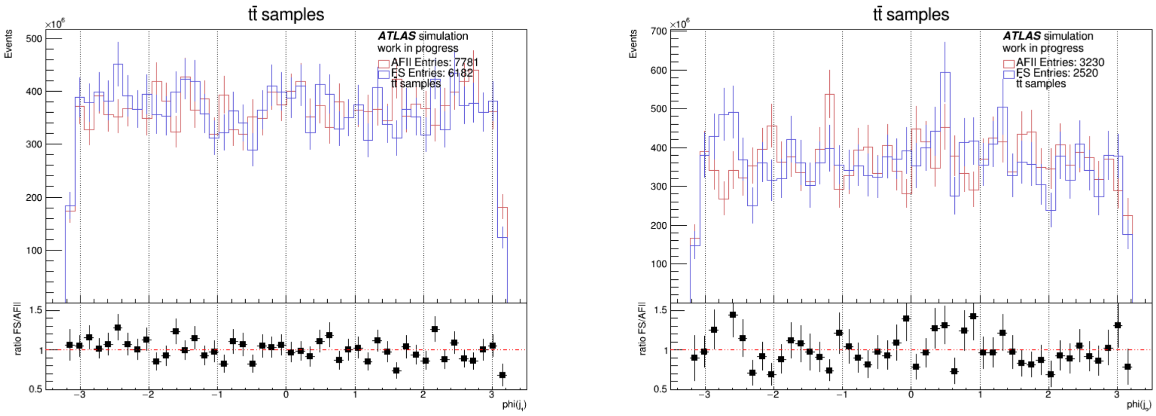

FIG. 6: ϕ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

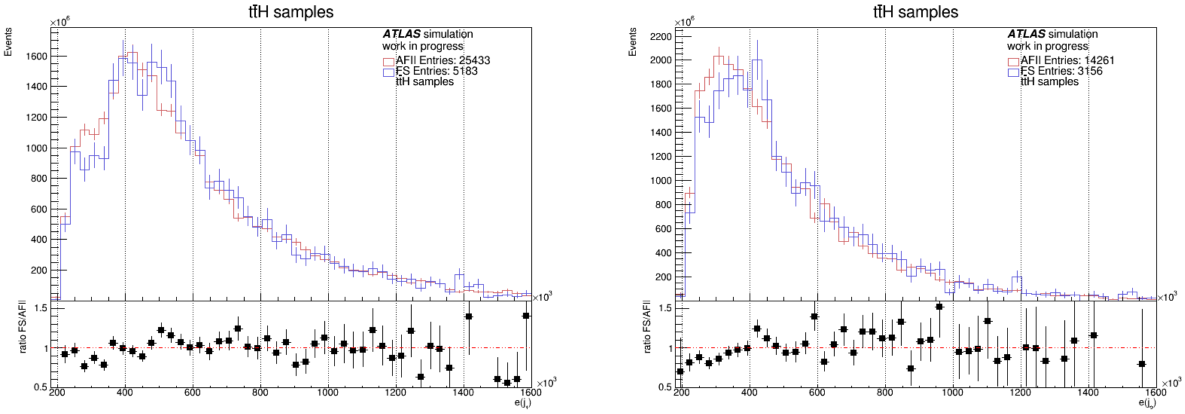

FIG. 7: e distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

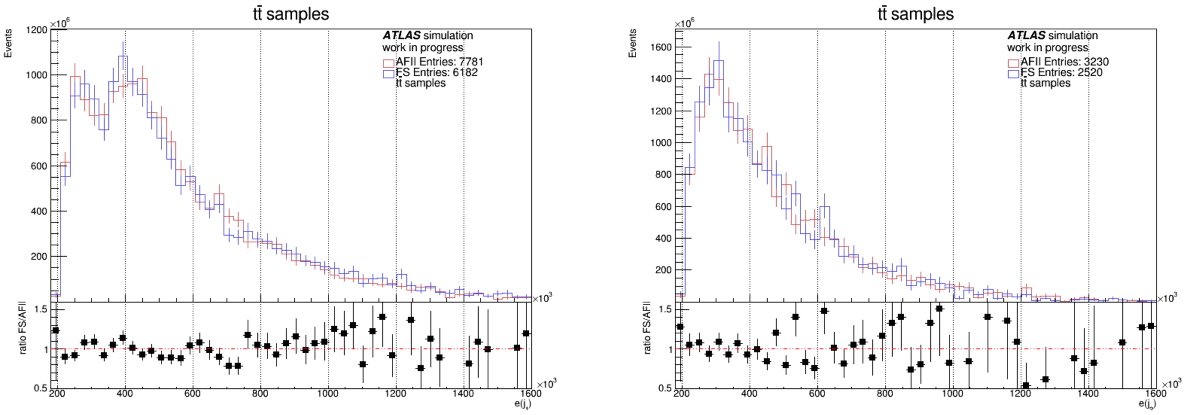

FIG. 8: e distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

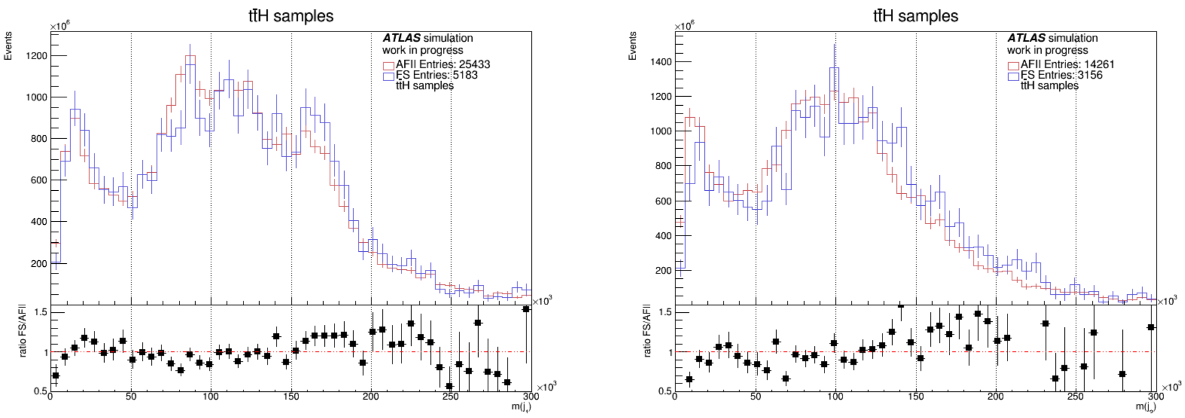

FIG. 9: m distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

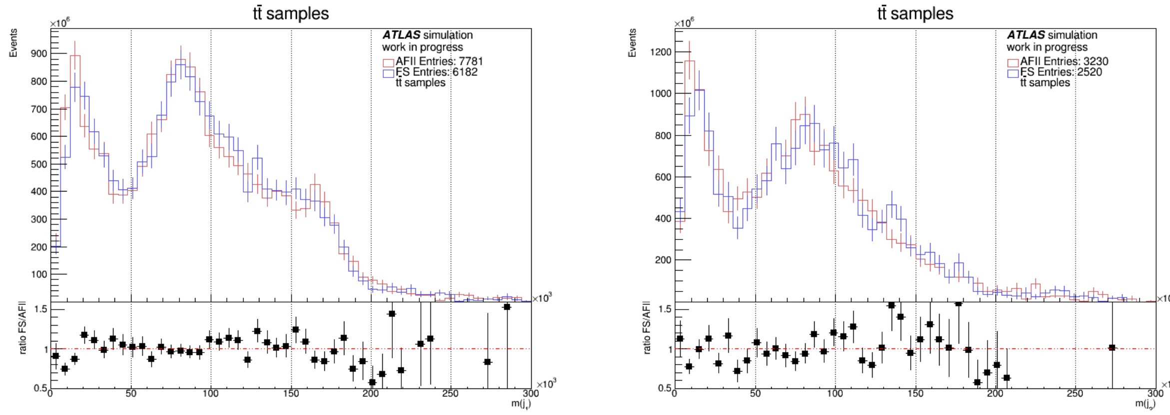

FIG. 10: m distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

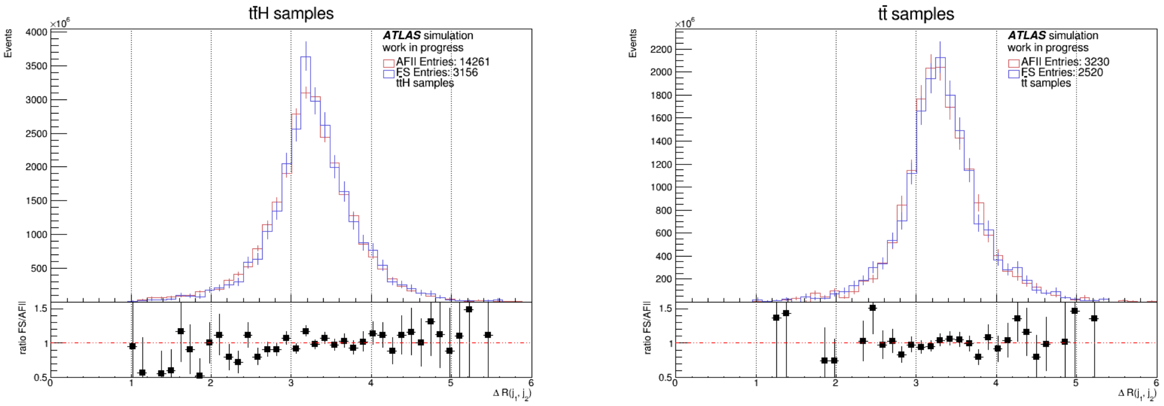

FIG. 11: ΔR distributions between leading and subleading large-R jets using $t\bar{t}H$ (left) and $t\bar{t}$ (right) samples generated with AF-II and FS detector simulations

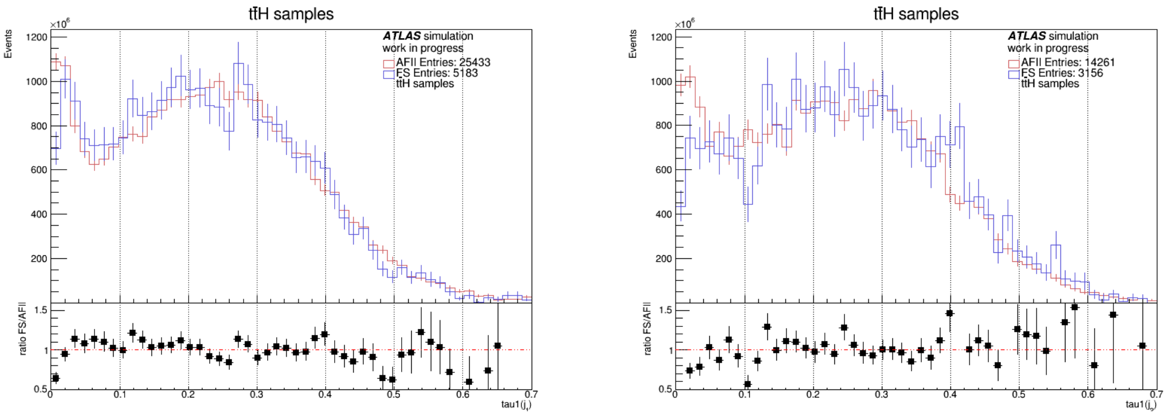

FIG. 12: τ_1^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

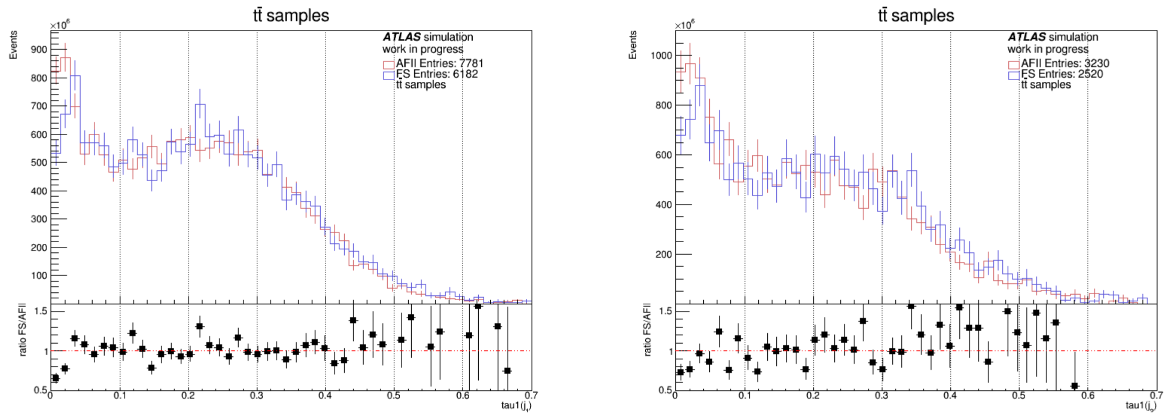

FIG. 13: τ_1^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

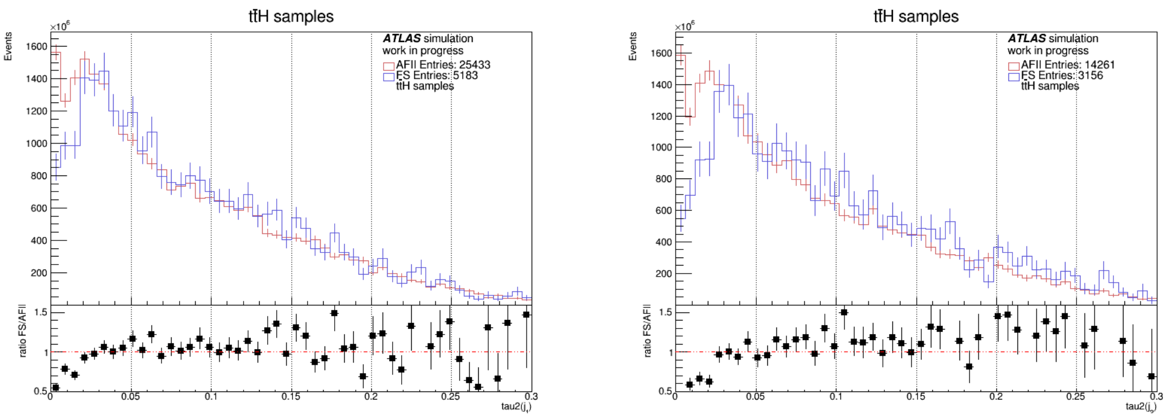

FIG. 14: τ_2^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

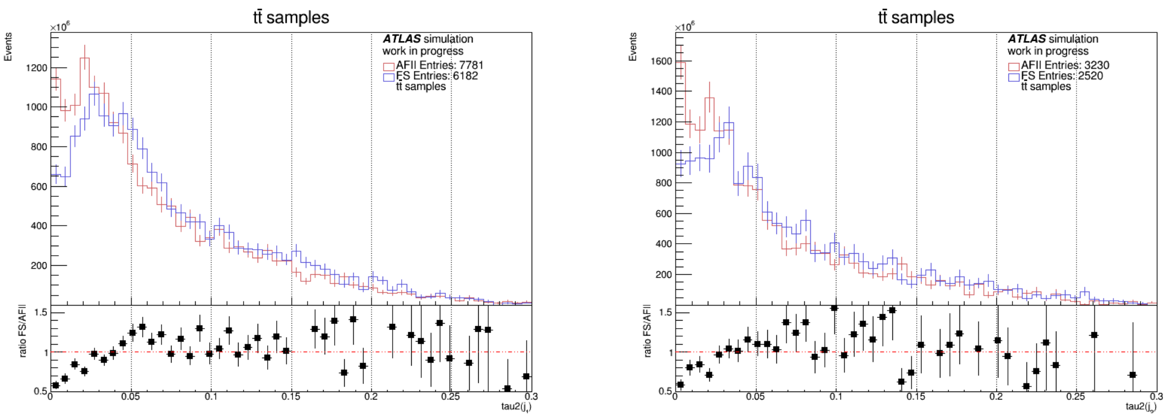

FIG. 15: τ_2^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

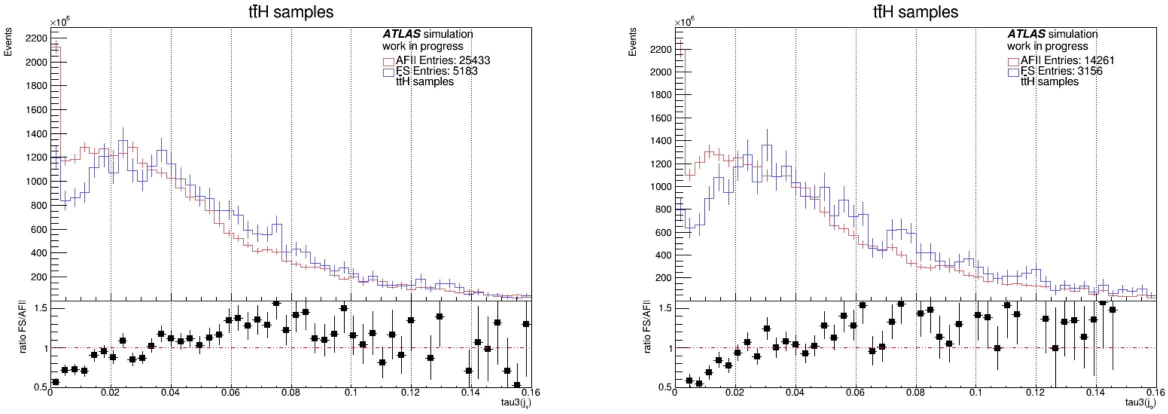

FIG. 16: τ_3^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

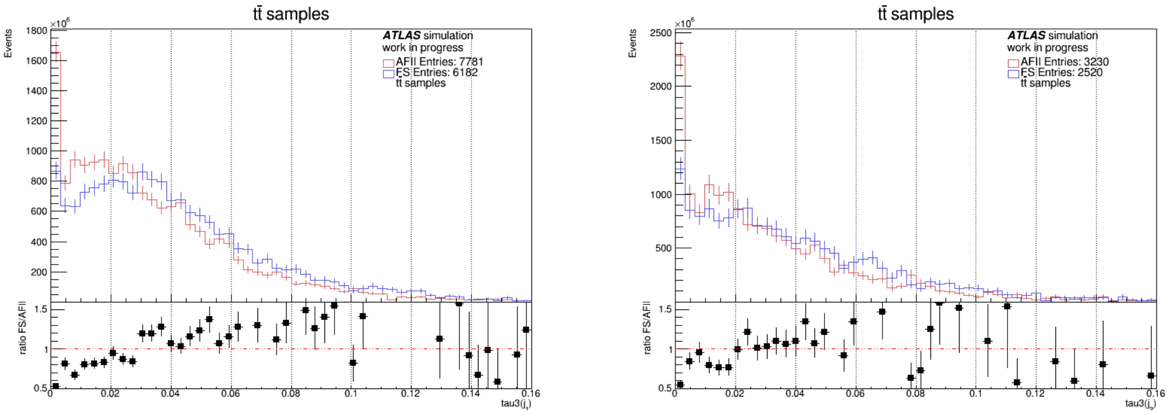

FIG. 17: τ_3^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

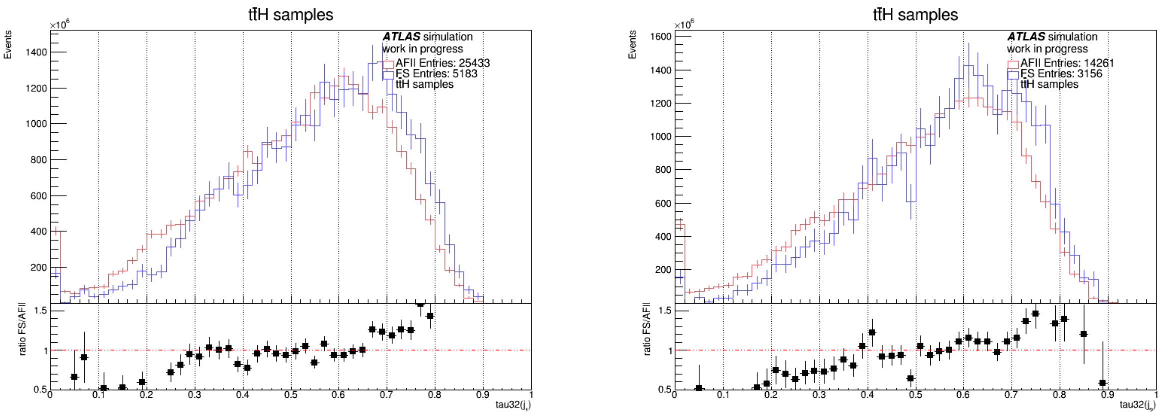

FIG. 18: τ_{32}^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

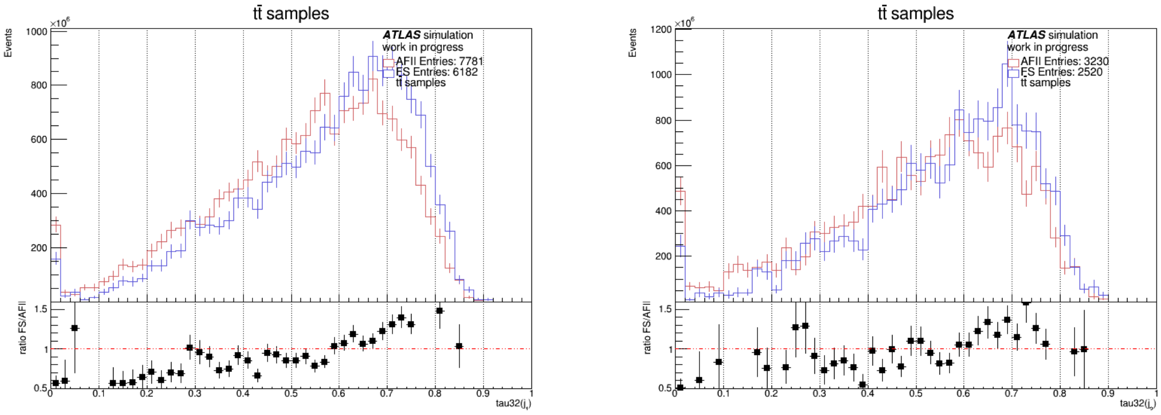

FIG. 19: τ_{32}^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

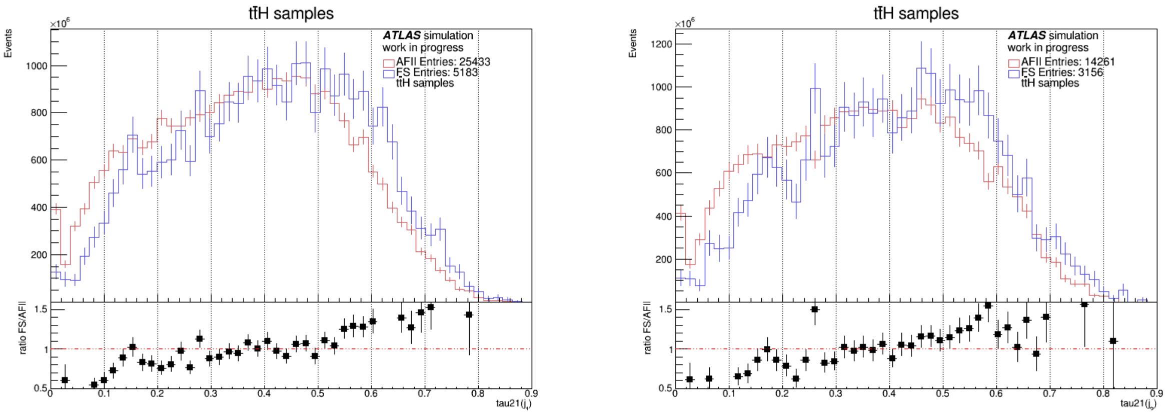

FIG. 20: τ_{21}^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

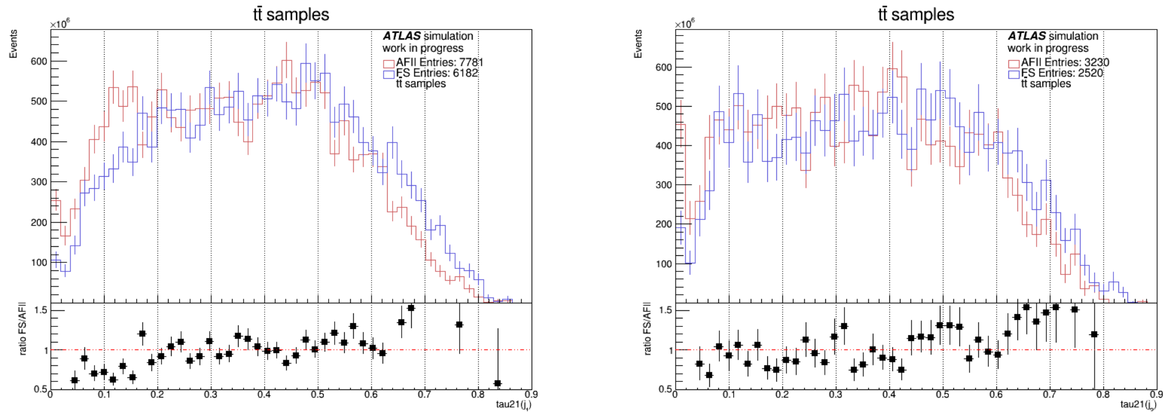

FIG. 21: τ_{21}^{wta} distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

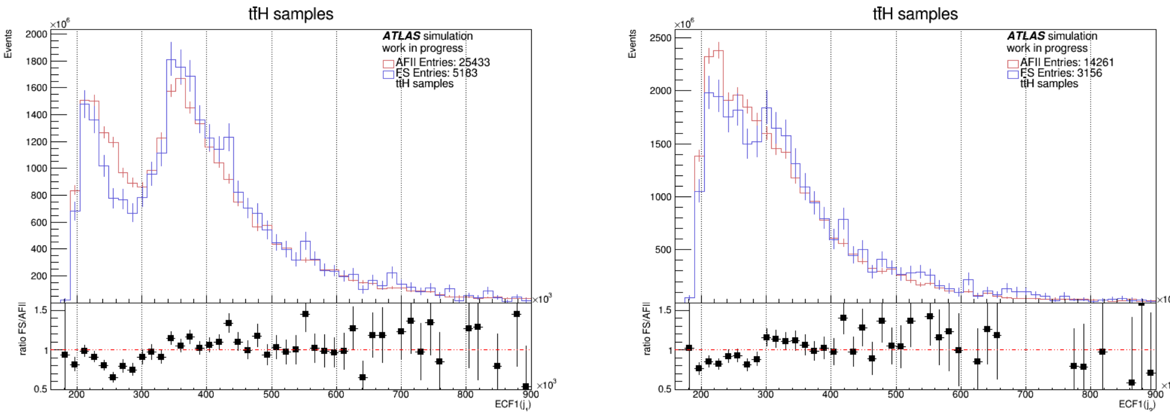

FIG. 22: ECF_1 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

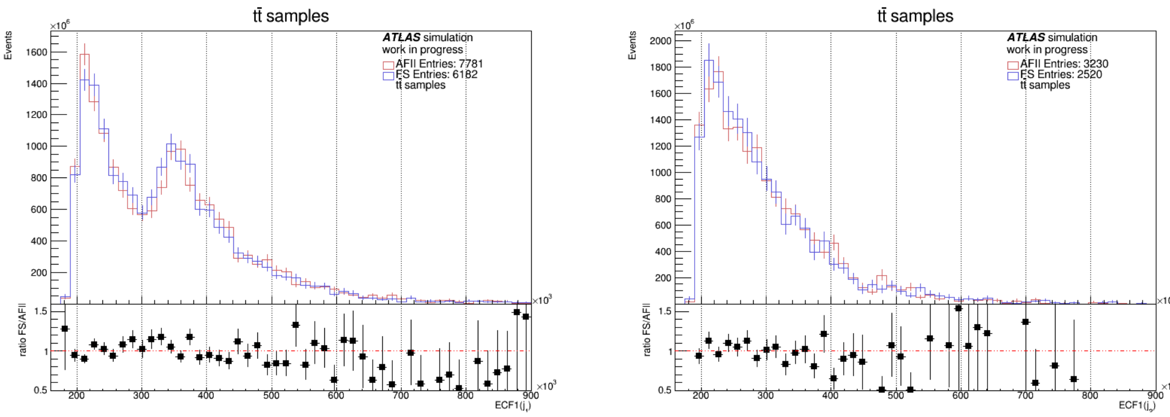

FIG. 23: ECF_1 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

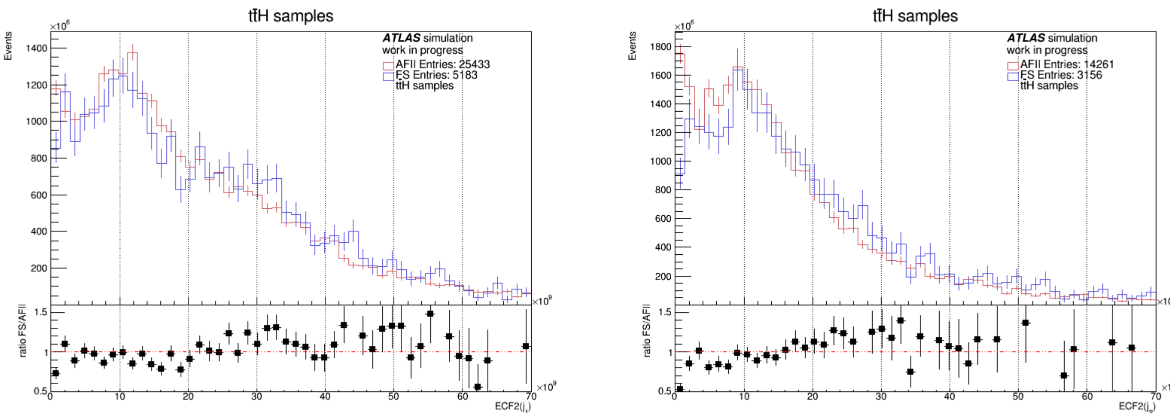

FIG. 24: ECF_2 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

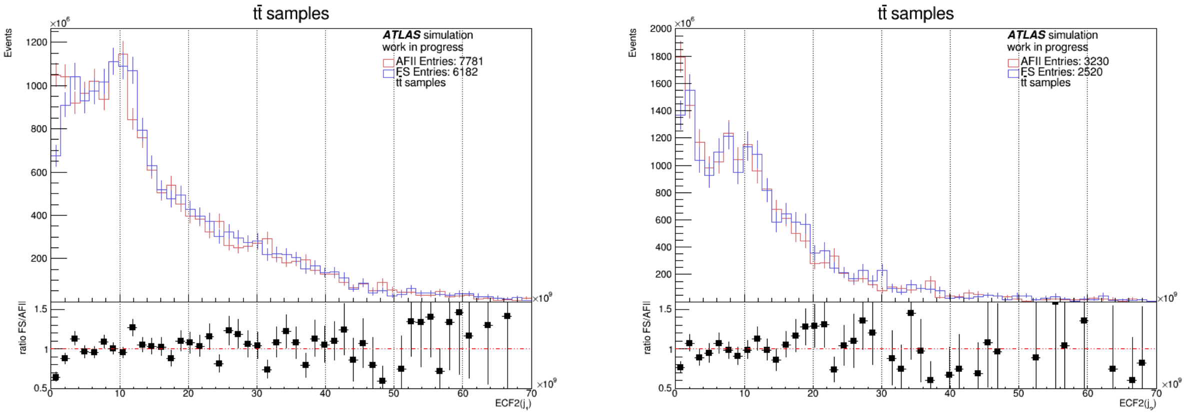

FIG. 25: ECF_2 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

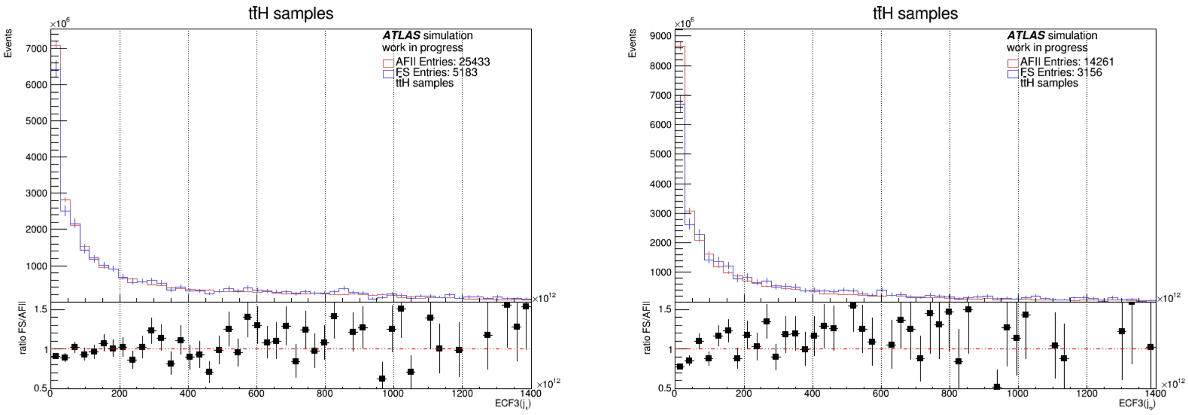

FIG. 26: ECF_3 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

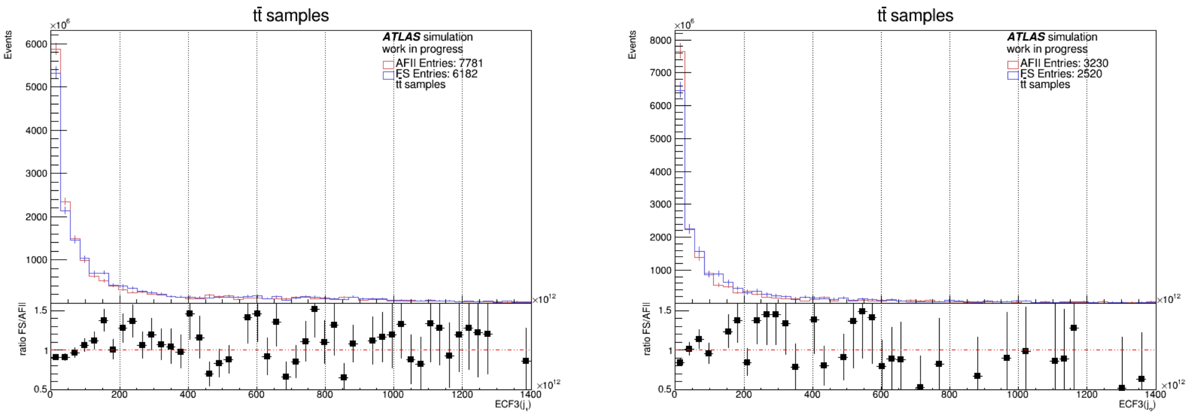

FIG. 27: ECF_3 distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

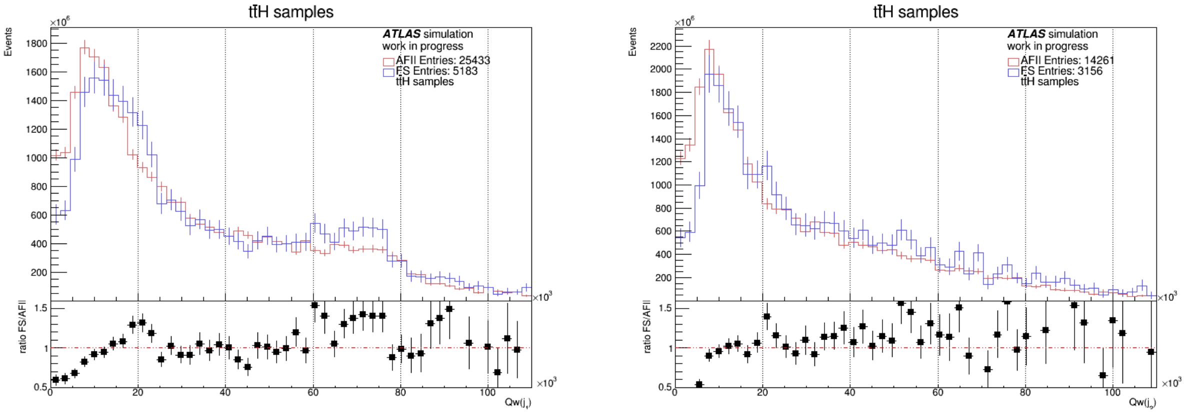

FIG. 28: Qw distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

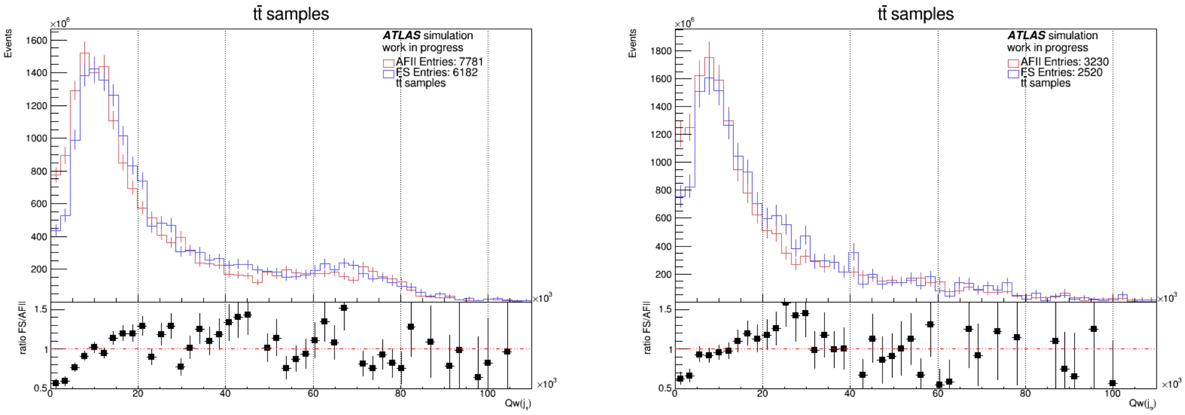

FIG. 29: Qw distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

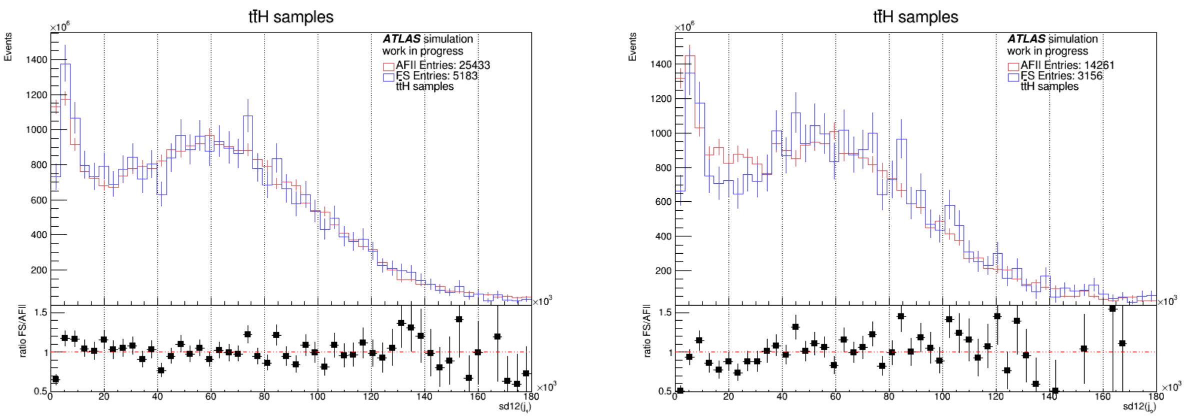

FIG. 30: $Split_{12}$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

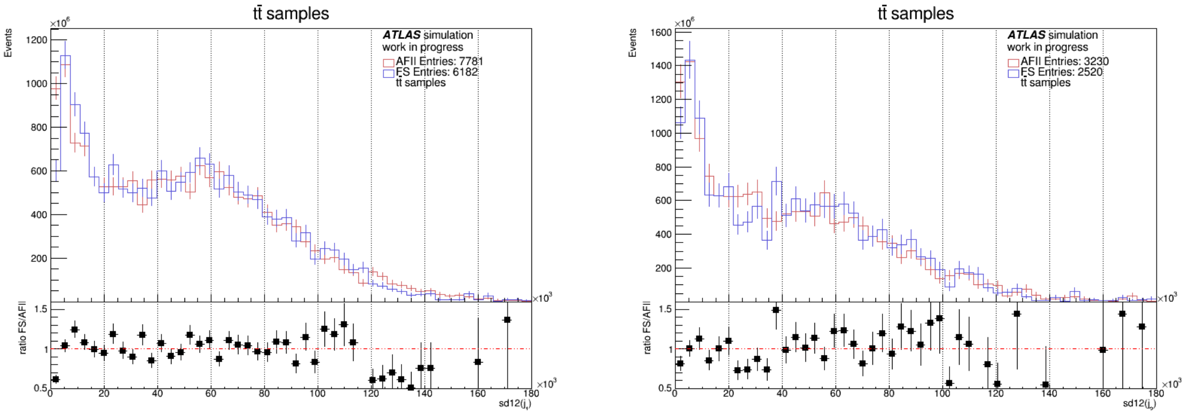

FIG. 31: $Split_{12}$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

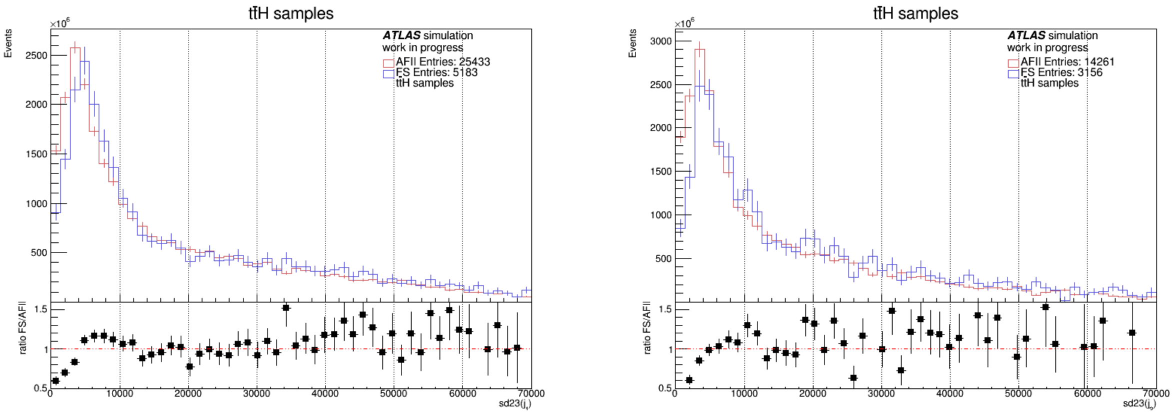

FIG. 32: $Split_{23}$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

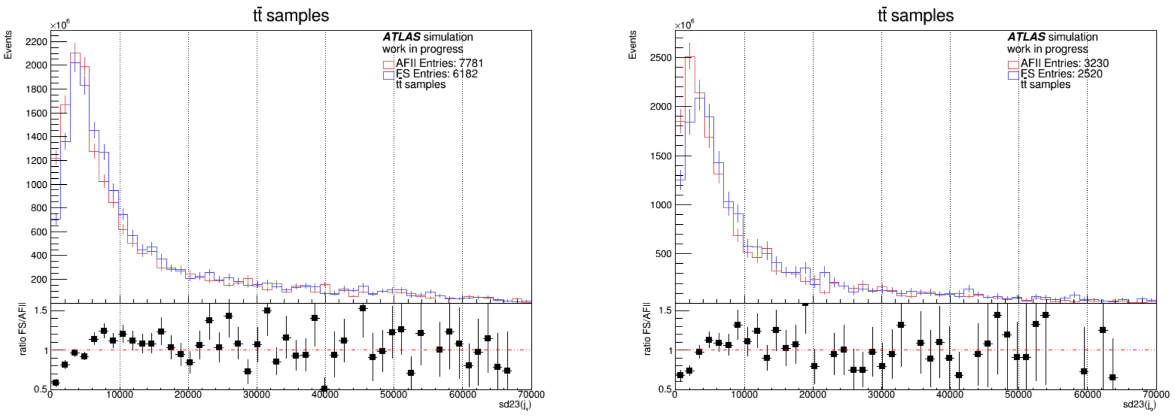

FIG. 33: $Split_{23}$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

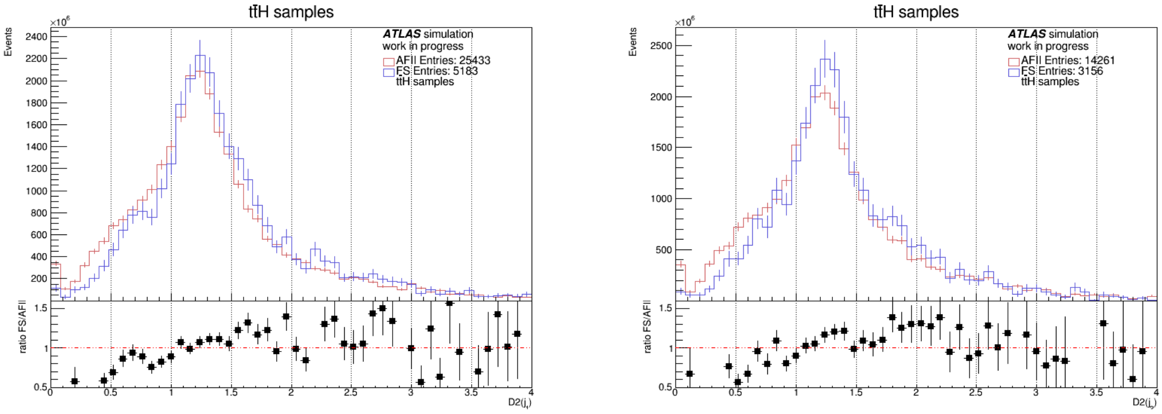

FIG. 34: $D2$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

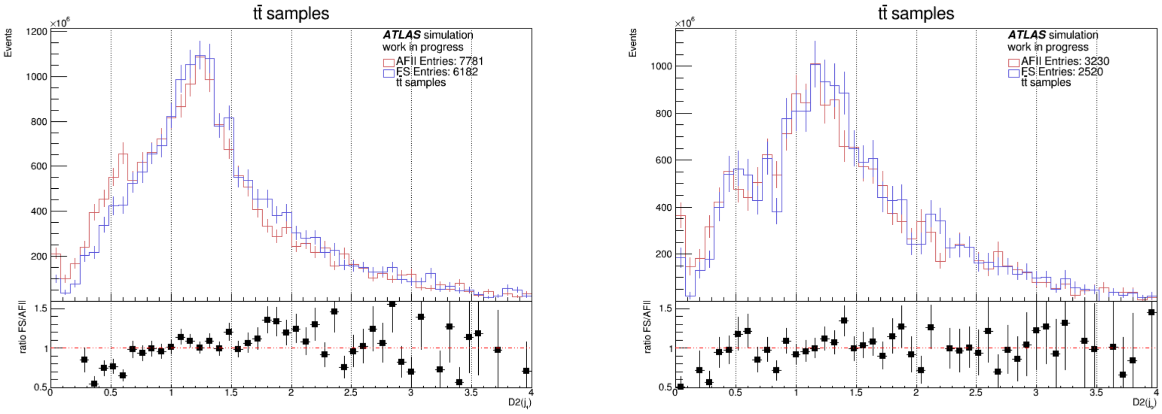

FIG. 35: $D2$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

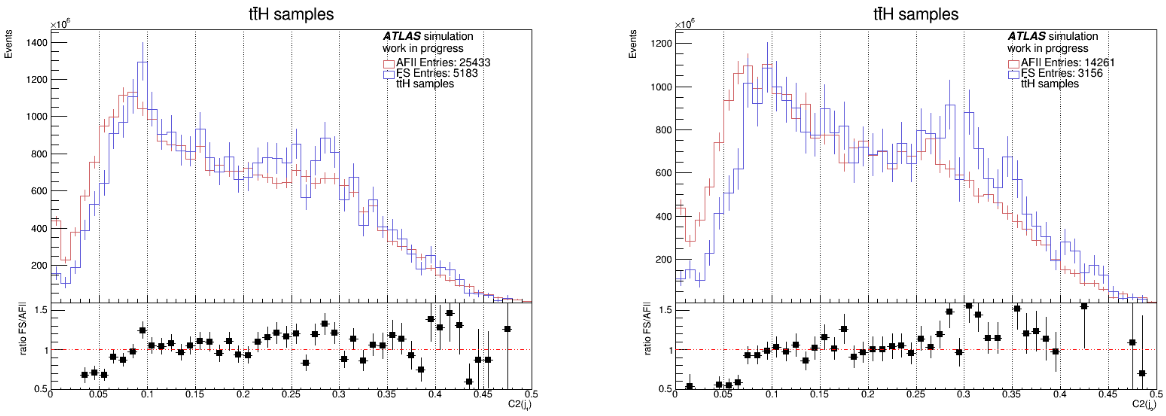

FIG. 36: $C2$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

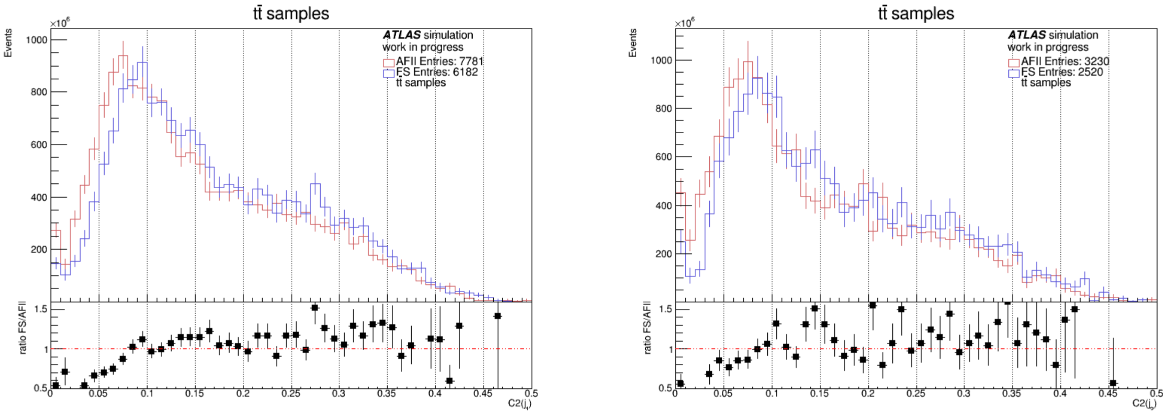

FIG. 37: $C2$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations

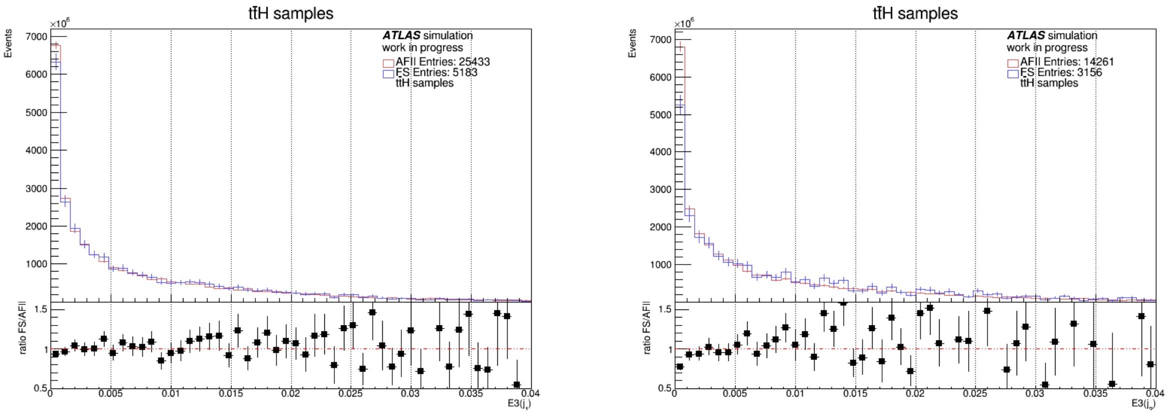

FIG. 38: $E3$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}H$ samples generated with AF-II and FS detector simulations

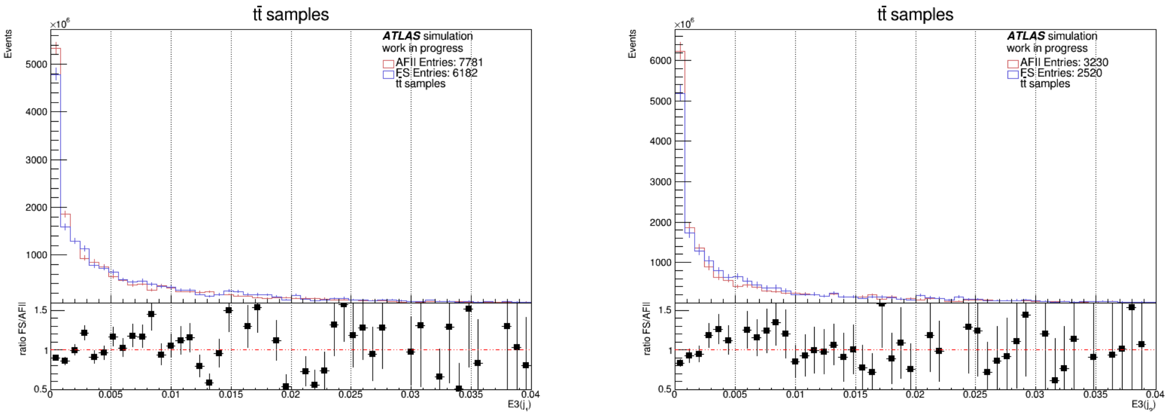

FIG. 39: $E3$ distributions for the leading (left) and subleading (right) large-R jets using $t\bar{t}$ samples generated with AF-II and FS detector simulations