Jet Radius Dependence of Systematic Uncertainties in Semileptonic Top Pair Production

Tobias Neuhoff[1]*
Supervisors: Tancredi Carli[2], Anastasia Kotsokechagia[2]

[1]University of Bonn
[2]ATLAS Collaboration, CERN

Abstract

The dependence of modelling uncertainties on the jet radius parameter has been studied in semileptonic $t\bar{t}$ events at particle level using Rivet. Variations of the Pythia *A14 tune, including initial- and final-state radiation variations, as well as colour reconnection and recoil models, were examined. Final-state radiation is found to be the dominant source of uncertainty for small jet radii, reaching the level of 10%, while initial-state radiation and scale variations on the cross section computation of the hard scattering have a much smaller impact of around 1%. The generator dependence, in particular the difference between* Pythia *and* Herwig*, remains one of the largest uncertainties in $t\bar{t}$ measurements, with effects of up to 20% for small jet radii. Studies of jet constituents and baryon production were performed as a first step towards understanding these differences. The results indicate that theoretical modelling uncertainties are strongly dependent on the chosen jet radius, with larger radii mitigating these uncertainties in jet calibration and top-quark measurements.*

Introduction

Accurate modelling of the W boson and the top-quark (t) is essential for precision measurements of top-quark properties, jet energy calibration, and the validation of Monte Carlo (MC) event generators used in ATLAS analyses. Semileptonic $t\bar{t}$ events provide an ideal laboratory: the presence of a high-p_T lepton allows for efficient triggering, the large $t\bar{t}$ production cross section ensures sufficient statistics to study precisely theoretical uncertainties. A schematic of the semileptonic $t\bar{t}$ process is shown in Figure 1.

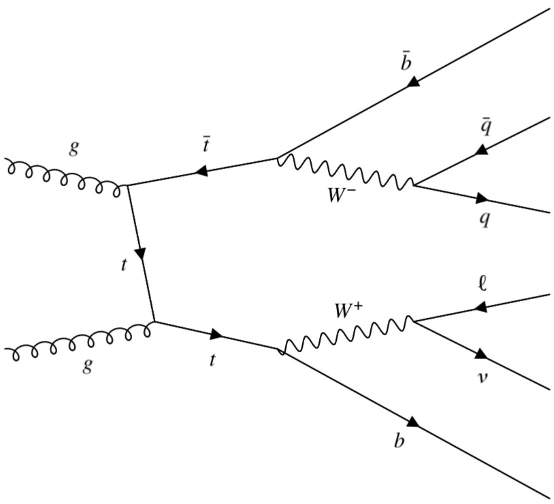

Figure 1: *Feynman diagram of a semileptonic $t\bar{t}$ event.*

Jets, reconstructed from collimated sprays of hadrons, are sensitive to both perturbative and non-perturbative QCD effects, including initial- and final-state radiation (ISR and FSR), contributions from the underlying event (UE) and hadronization. These effects are modelled with phenomenological approaches in MC generators such as Pythia 8 [1], Herwig [2], and Sherpa [3]. Variations of parameters controlling ISR, FSR, UE, and colour reconnection (CR) are used to estimate systematic modelling uncertainties, which remain among the dominant theoretical uncertainties in many ATLAS measurements.

The present study investigates the impact of the jet radius R on these modelling uncertainties. The jet radius choice impacts the balance between capturing radiation from the hard parton and avoiding contamination from UE and nearby jets, and thereby affects the reconstruction of the W boson and top-quark masses. While the standard choice in ATLAS analyses has been $R = 0.4$ since the beginning of data taking, theoretical arguments suggest that a larger radius may reduce certain uncertainties [4]. However, no comprehensive study has revisited this choice in more than fifteen years.

This analysis is performed at particle level using a standard Rivet routine [5], considering semileptonic $t\bar{t}$ events generated with Powheg+Pythia 8 [6, 1] using the A14 tune [7], Herwig 7.1.3 and 7.2.1, and Sherpa 2.2.1 and 2.2.12. The key observables are the reconstructed W boson and top-quark masses. A wide range of generator variations, including renormalisation and factorisation scales (μ_R, μ_F), ISR/FSR, UE, different recoil schemes, and CR models, are

---
*CERN summer student report
Corresponding author: tobias.neuhoff@cern.ch
Published: September 29, 2025

considered to quantify the dependence of theoretical uncertainties on the jet radius.

The goal of this study is to provide guidance on the optimal choice of jet radius in semileptonic $t\bar{t}$ events, with the aim of improving both the precision of top-quark measurements and reducing systematics from MC event generators used across ATLAS analyses.

Analysis Setup

Jets are clustered using the anti-k_t algorithm [8] with jet radius parameters $R \in \{0.2, 0.3, 0.4, 0.45, 0.5, 0.55, 0.6, 0.7, 0.8, 0.9, 1.0\}$. Events are required to contain exactly one charged lepton (electron or muon) with $p_T > 30$ GeV and $|\eta| < 4.2$. To enhance the contribution from semileptonic $t\bar{t}$ decays, a missing transverse momentum requirement of $E_T^{miss} > 30$ GeV is imposed. Jets are reconstructed excluding the identified lepton and must have $p_T > 30$ GeV. Jets overlapping with the selected lepton within $\Delta R < 0.3$ are removed. Only events with at least four jets are retained.

Inclusive jet kinematics, such as jet multiplicity and transverse-momentum spectra, are studied before applying further top-quark reconstruction requirements. Jets are then categorised using b-tagging: exactly two jets are required to be b-tagged, while at least two further jets must be untagged (light jets).

The $t\bar{t}$ system is reconstructed using the two b-jets, the light-jet pair closest to the W boson mass, the selected lepton, and a neutrino candidate inferred from E_T^{miss} (Figure 2). The neutrino longitudinal momentum is determined by solving the quadratic constraint from the on-shell $W \to \ell\nu$ mass, following the standard reconstruction procedure employed by the ATLAS top-quark group. The hadronically decaying W boson is reconstructed from the light-jet pair whose invariant mass is closest to the nominal W boson mass. To suppress wrong pairings, events are vetoed if the reconstructed hadronic W boson mass lies outside $75 < m_W < 85$ GeV[1]. Top-quark candidates are formed by combining the W boson with each of the two b-jets in turn, since the correct assignment cannot be unambiguously identified. Both combinations are included in the distributions of reconstructed top-quark mass and hadronic-top transverse momentum. The leptonic top is reconstructed by combining the lepton, neutrino, and each b-jet in turn. The transverse momentum of the full $t\bar{t}$ system is also recorded.

Angular correlations between the highest-p_T b-jet and other objects can be studied with distributions in ΔR, $\Delta\eta$, and $\Delta\phi$ with respect to the second b-jet, the leading and subleading light jets, and the reconstructed W boson[2].

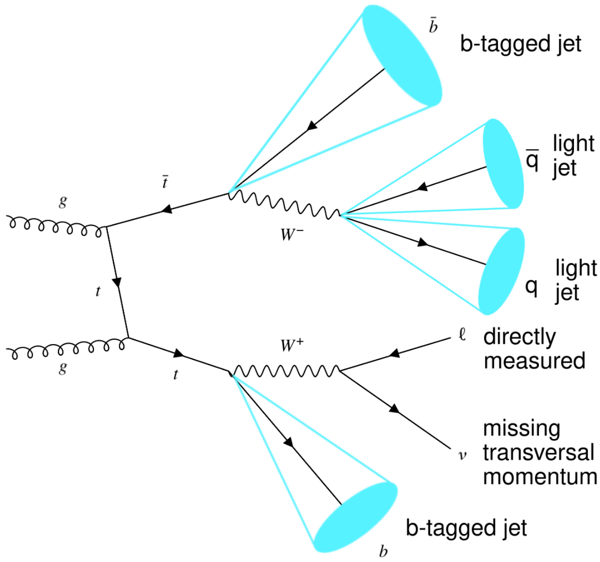

Figure 2: *Reconstruction of a semileptonic $t\bar{t}$ event using anti-k_t jets.*

Results and Discussion

Pythia A14 Variations

The A14 tune variations [7] in $t\bar{t}$ production, as prescribed by the ATLAS top-quark group [9], are first considered.

Figure 3 presents the reconstructed W boson mass under variations of the initial-state radiation (ISR) parameters. The corresponding $t\bar{t}$ top-quark mass distribution is shown in Figure 4. The ISR uncertainty around the W boson mass is found to be below the percent level and exhibits a jet radius dependence. For the top-quark mass, the effect remains below 2% for all radii.

The final-state radiation (FSR) contribution is varied by increasing or decreasing the renormalisation scale used in the FSR parton shower. The impact of FSR variations is displayed in Figures 5 and 6 and decreases substantially with larger jet radii: for $R = 0.4$ the modification reaches approximately 6% in the W boson mass, while for $R = 0.6$ it is reduced to about 3%. A similar behaviour is observed in the top-quark mass.

Underlying-event (UE) variations are shown in Figures 7 and 8. At $R = 0.4$ the maximal difference reaches approximately 6% in the W boson mass.

[1] Figure 23 demonstrates the difference between applying and not applying the W mass cuts.

[2] Here, $\Delta\eta$ and $\Delta\phi$ denote the differences in pseudorapidity and azimuthal angle between two objects, respectively. The variable $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is a measure of the angular separation in the detector.

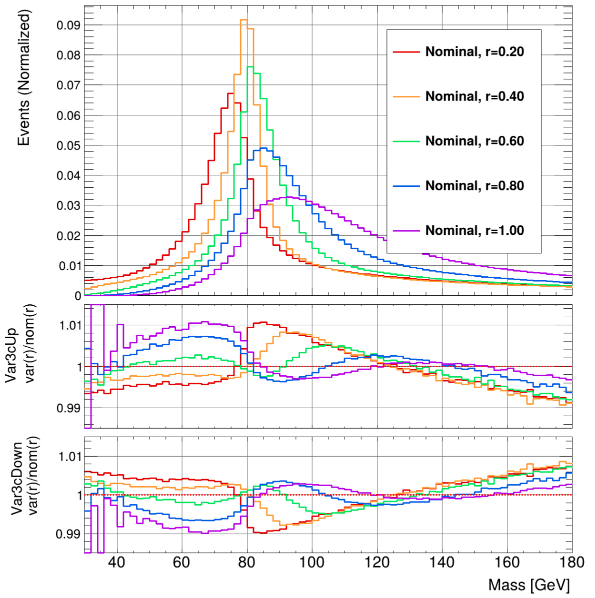

Figure 3: *Reconstructed W boson mass for different jet radii under ISR variations.*

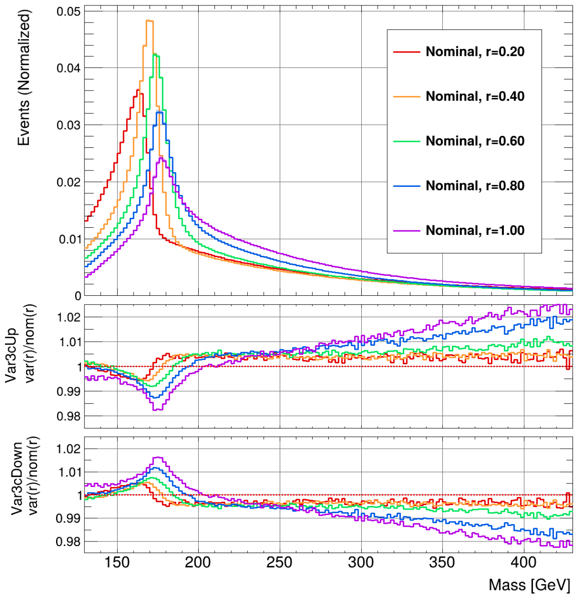

Figure 4: *Reconstructed top-quark mass for different jet radii under ISR variations.*

Slightly higher effects are observed for the top-quark mass.

The variations of the renormalisation and factorisation scales of the hard scattering yield differences of less than 2% (Figure 25).

Model Differences

The impact of different colour reconnection (CR) models implemented in PYTHIA 8 is investigated, namely the MPI-based model, the QCD-based model, and the gluon-move model [10, 11]. In addition, the `RecoilToTop` option, which redistributes recoil between the W boson system and the top quarks, is studied.

The MPI-based CR model (Figures 10, 11) shows to have an increasing difference to the nominal for higher jet radii. At $R = 0.4$ the deviation around the W boson mass corresponds to about 3%.

The QCD-based CR model (Figures 12, 13) exhibits a minor effect. At $R = 0.4$ the deviation in the W boson mass corresponds to about 2%. The difference remains below 5% for all radii in W boson and top-quark mass.

The gluon-move CR model (Figures 14, 15) shows to have minor impact as well. At $R = 0.4$ the deviation corresponds to about 2% in the W boson mass. The difference remains below 5% for all radii in W boson and top-quark mass.

The `RecoilToTop` model (Figures 16, 17) shows a difference of about 1% at $R = 0.4$ in the W boson mass. For the top-quark mass, the effect is larger. In the `RecoilToTop` model, gluon emissions in the top-quark decay are recoiled coherently onto the top quark itself (rather than on color partners), thereby modifying out-of-cone radiation and partially mitigating biases in jet kinematics [12, 13].

Generator Differences

Differences between MC generators, in particular between PYTHIA and HERWIG, represent one of the largest theoretical uncertainties in top-quark measurements and jet calibration. Despite that both generators are tuned to similar datasets, significant discrepancies remain.

Figure 18 presents the W boson mass distribution in PYTHIA and HERWIG for different jet radius choices. The difference is characterised by both a peak shift to higher values and a broadening of the distribution, both of which decrease with increasing radius. At $R = 0.4$ the difference reaches about 20%. For the top-quark mass, similar effects are observed, though less pronounced.

To further investigate, two versions of SHERPA (2.2.1 and 2.2.12, see [3, 14]) are compared to PYTHIA and HERWIG [15, 16] for a jet radius of $R = 0.4$ and $R = 0.6$. The newer Sherpa version is tuned to better reproduce LEP data and in particular baryon fractions [14]. However, both Sherpa versions yield very similar results for the W boson and top-quark masses (Figures 20, 21). In contrast to HERWIG, which shifts the peak to the right, Sherpa shifts the peak to the left. At $R = 0.4$, Sherpa differs from Pythia by about 10%. Moreover, the Sherpa-Pythia difference increases with jet radius, while the Herwig-Pythia difference decreases.

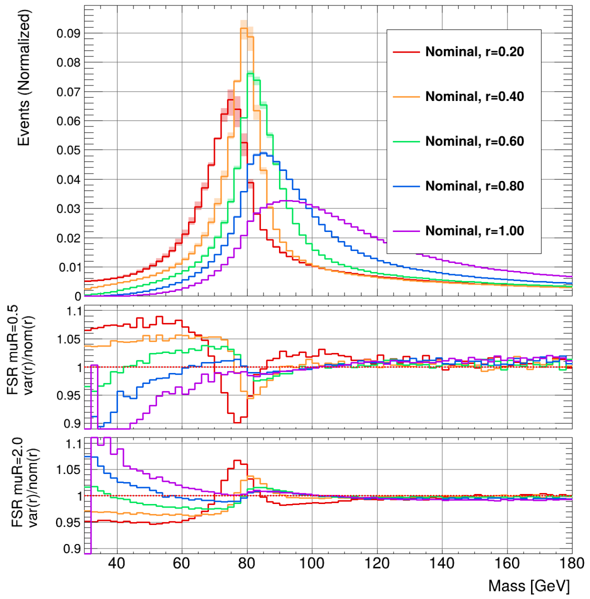

Figure 5: *Reconstructed W boson mass for different jet radii under FSR variations.*

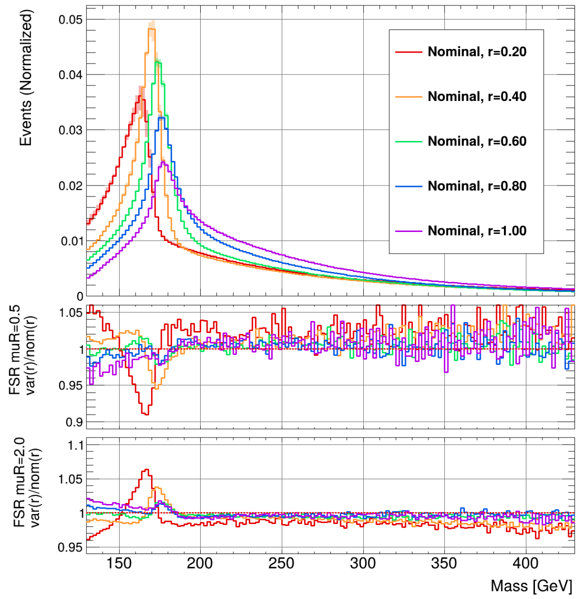

Figure 6: *Reconstructed top-quark mass for different jet radii under FSR variations.*

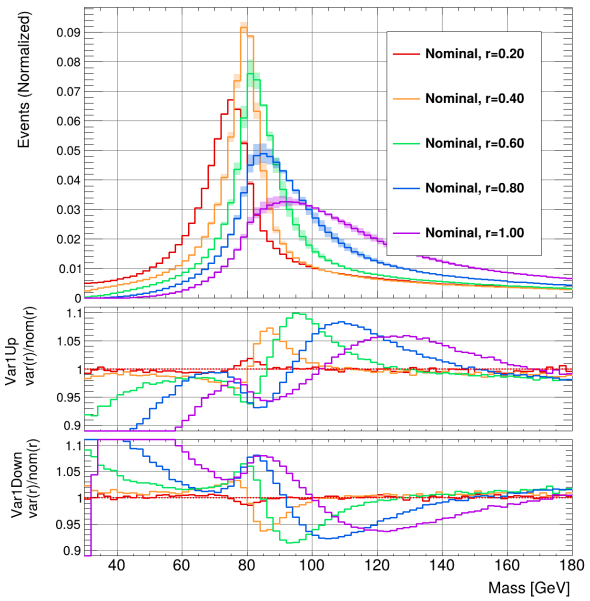

Figure 7: *Reconstructed W boson mass for different jet radii under UE variations.*

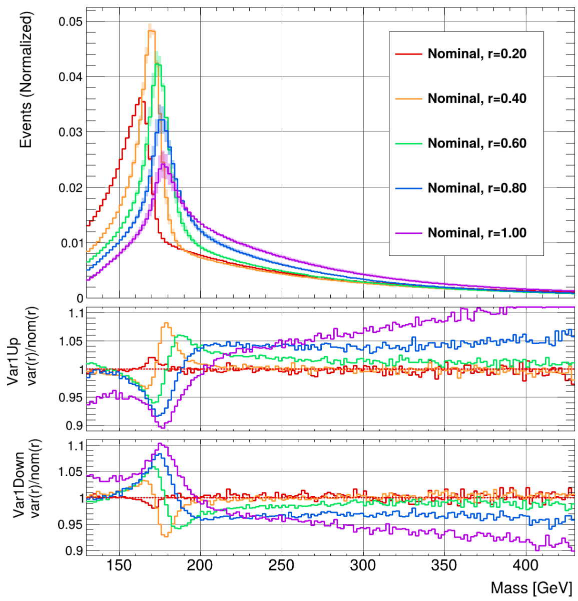

Figure 8: *Reconstructed top-quark mass for different jet radii under UE variations.*

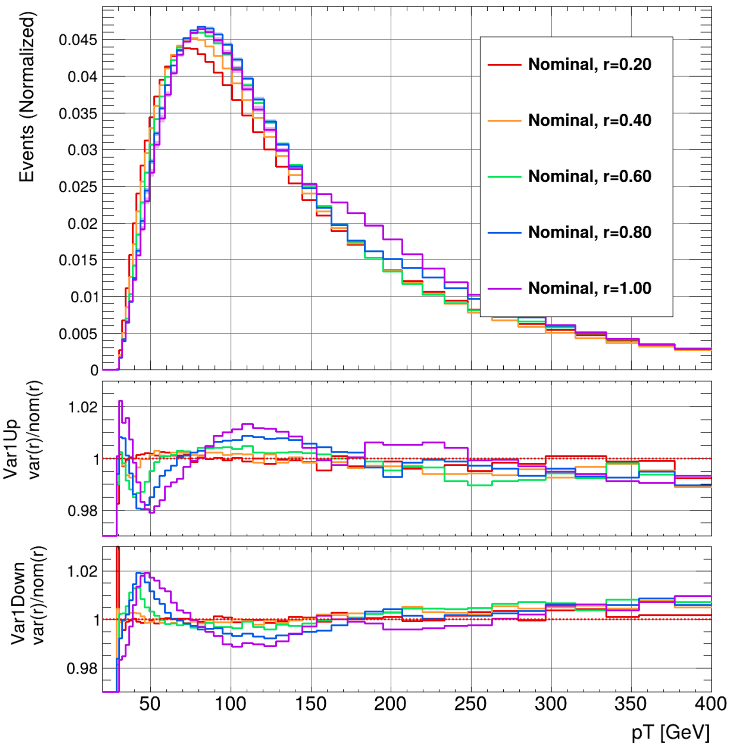

Figure 9: *Transverse momentum of the leading non-b-tagged jet for different jet radii under UE variations.*

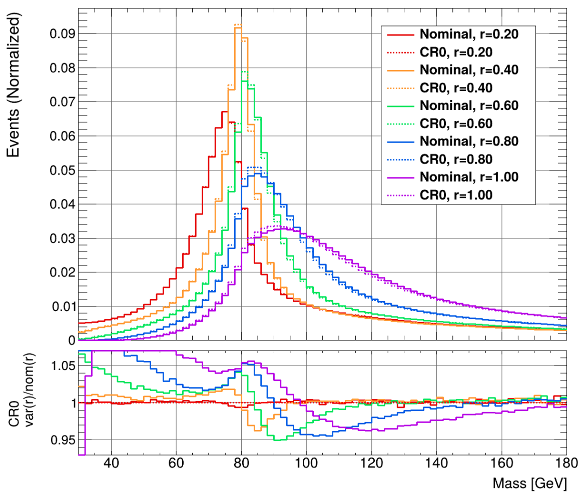

Figure 10: *Reconstructed W boson mass for different jet radii with the MPI-based CR model.*

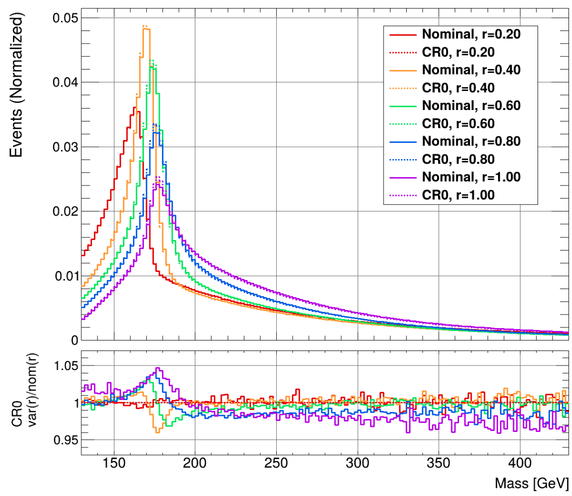

Figure 11: *Reconstructed top-quark mass for different jet radii with the MPI-based CR model.*

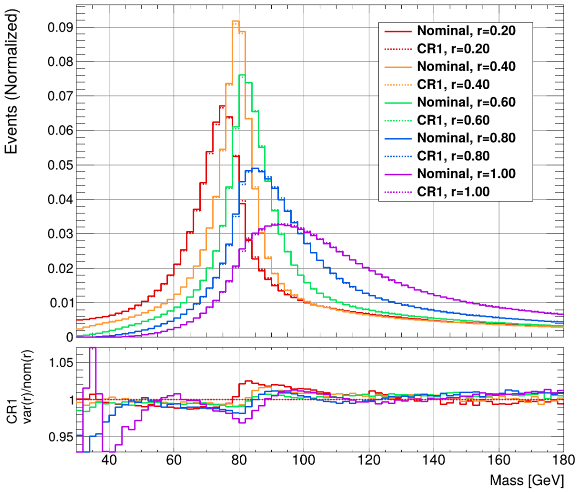

Figure 12: *Reconstructed W boson mass for different jet radii with the QCD-based CR model.*

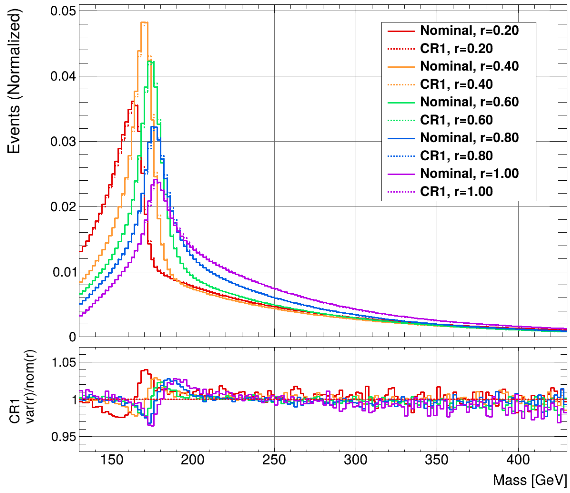

Figure 13: *Reconstructed top-quark mass for different jet radii with the QCD-based CR model.*

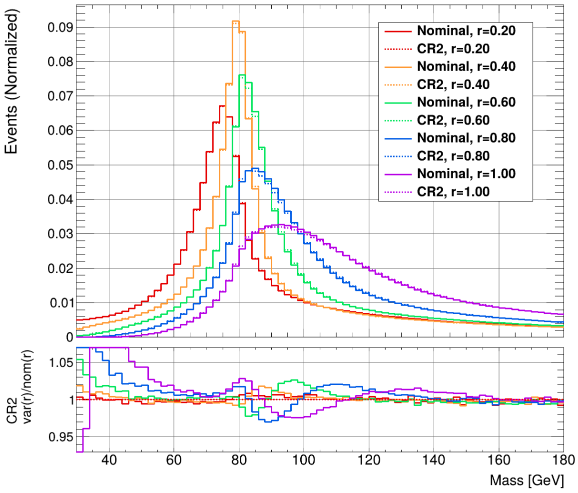

Figure 14: *Reconstructed W boson mass for different jet radii with the gluon-move CR model.*

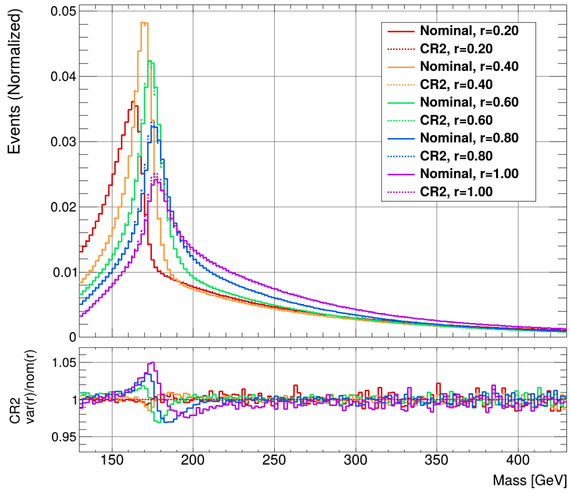

Figure 15: *Reconstructed top-quark mass for different jet radii with the gluon-move CR model.*

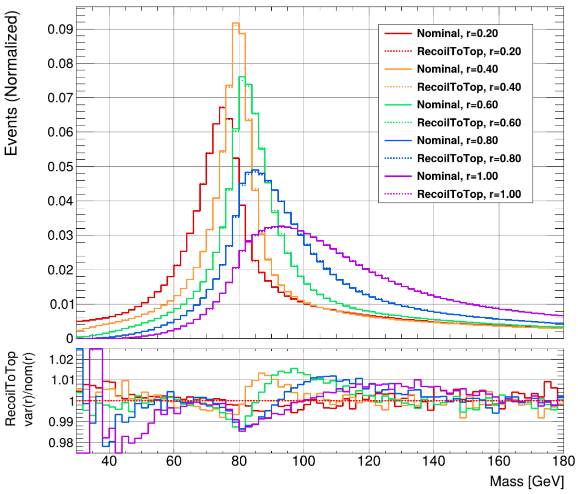

Figure 16: *Reconstructed W boson mass for different jet radii with the* `RecoilToTop` *option in* PYTHIA.

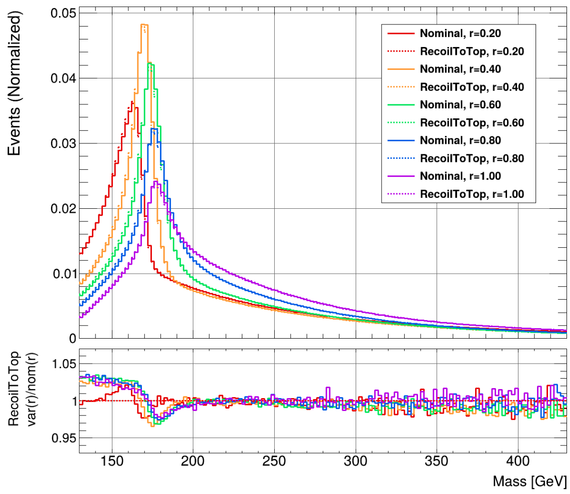

Figure 17: *Reconstructed top-quark mass for different jet radii with the* `RecoilToTop` *option in* PYTHIA.

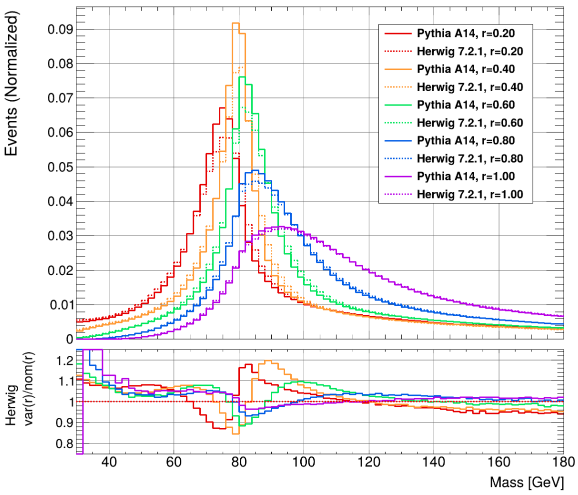

Figure 18: *Reconstructed W boson mass for different jet radii:* Pythia *vs* Herwig.

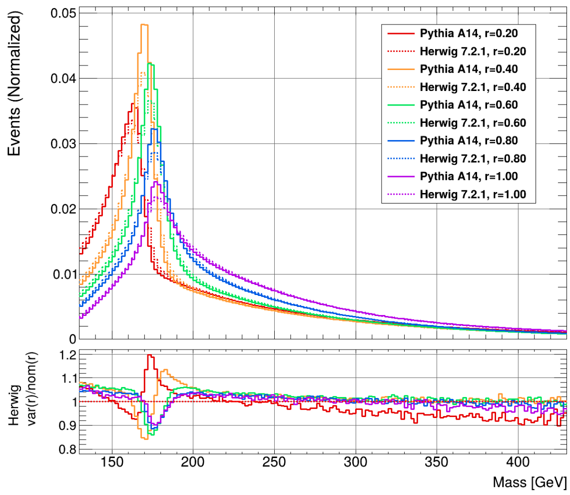

Figure 19: *Reconstructed top-quark mass for different jet radii:* Pythia *vs* Herwig.

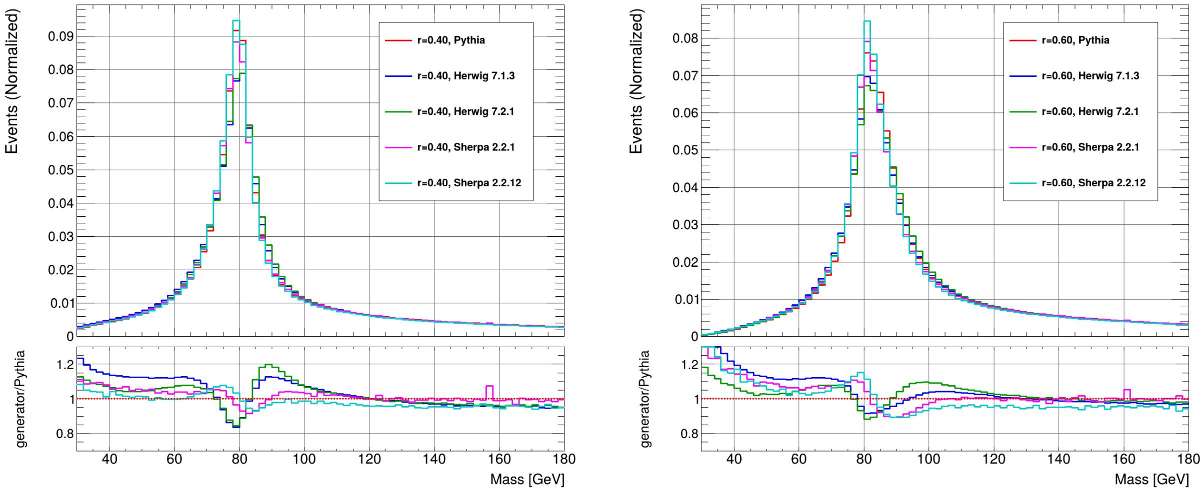

Figure 20: *Reconstructed W boson mass for different jet radii ($R = 0.4$, $R = 0.6$): generator comparison.*

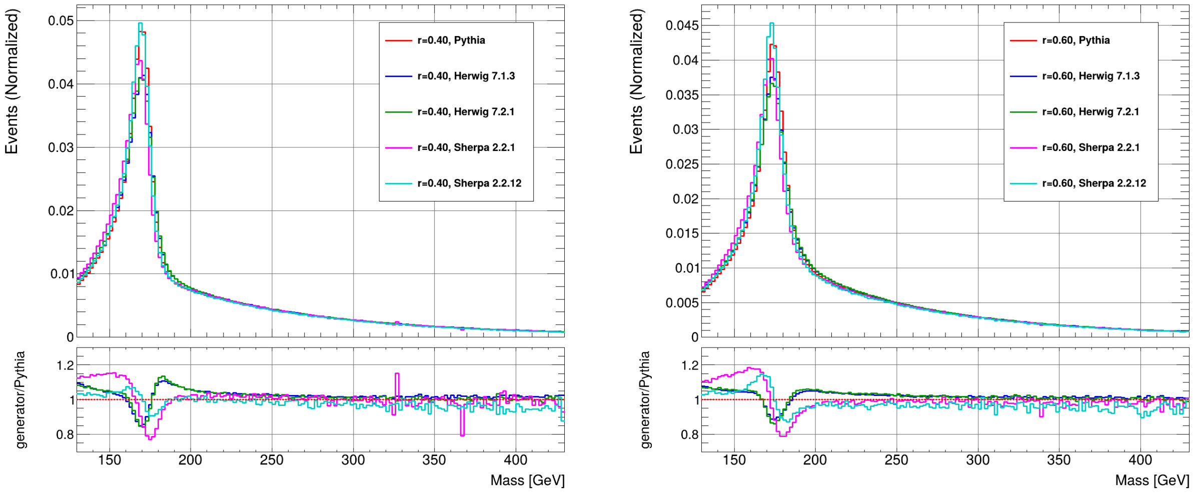

Figure 21: *Reconstructed top-quark mass for different jet radii ($R = 0.4$, $R = 0.6$): generator comparison.*

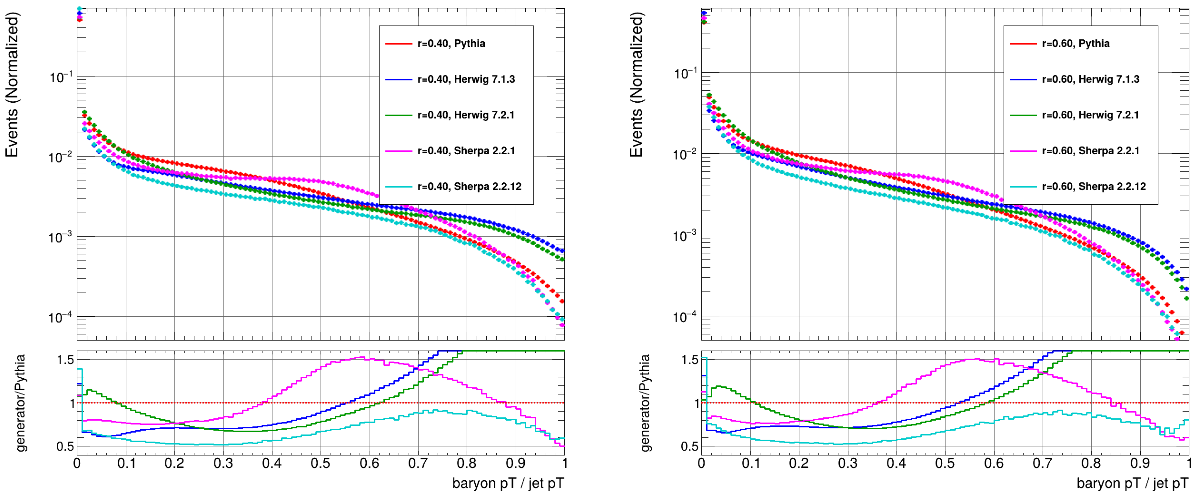

Figure 22: *Baryon fraction in the hadronic W boson jets for different generators and jet radii.*

A known difference between PYTHIA and HERWIG is the baryon fraction in jets originating from the hadronic W boson decay [17, 18, 19]. This motivates a more detailed study of the baryon fraction in the light jets of the reconstructed W boson.

Conclusion

The dependence of modelling uncertainties on the jet radius parameter has been investigated in semileptonic $t\bar{t}$ events at particle level.

Final-state radiation (FSR) variations are found to have the largest impact, reaching the level of 10% for small jet radii. Increasing the jet radius significantly reduces the FSR uncertainty, consistent with expectations. Initial-state radiation (ISR) and scale variations, by contrast, have only minor effects, typically around 1%.

For variations of colour reconnection and recoil models in PYTHIA, the behaviour with respect to the jet radius is less pronounced. The MPI-based model shows a significant effect, while the other models show a neglegible impact. Further dedicated studies are required to fully understand the observed differences.

The discrepancy between PYTHIA and HERWIG remains one of the dominant sources of theoretical uncertainty in $t\bar{t}$ measurements, with differences reaching up to 22% at small jet radii. This translates to a shift in the top-quark and W boson masses. A first exploration of jet constituents and their kinematic properties was performed, but no definitive explanation has yet been established.

Overall, these observations point towards the use of larger jet radii as a means to reduce the dominant FSR and generator-related uncertainties. The increase in UE-related effects requires further scrutiny, but appears less critical in experimental analyses.

References

[1] T. Sjostrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph]. URL: https://arxiv.org/abs/1410.3012.

[2] J. Bellm et al. "Herwig 7.0 / Herwig++ 3.0 Release Note". In: *Eur. Phys. J. C* 76 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph]. URL: https://arxiv.org/abs/1512.01178.

[3] E. Bothmann et al. "Event generation with SHERPA 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph]. URL: https://arxiv.org/abs/1905.09127.

[4] Mrinal Dasgupta, Lorenzo Magnea and Gavin P Salam. "Non-perturbative QCD effects in jets at hadron colliders". In: *Journal of High Energy Physics* 2008.02 (Feb. 2008), pp. 055–055. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/02/055. URL: http://dx.doi.org/10.1088/1126-6708/2008/02/055.

[5] A. Buckley et al. "Rivet user manual". In: *Comput. Phys. Commun.* 184 (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph]. URL: https://arxiv.org/abs/1003.0694.

[6] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 2010 (), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[7] *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. ATL-PHYS-PUB-2014-021. CERN, Dec. 2014. URL: https://cds.cern.ch/record/1966419.

[8] M. Cacciari, G. P. Salam and G. Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph]. URL: https://arxiv.org/abs/0802.1189.

[9] *Calibration of the jet energy scale and resolution of small-radius jets using semileptonic $t\bar{t}$ events with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2024-939. CERN, Oct. 2024. URL: https://cds.cern.ch/record/2915226.

[10] Ilkka Helenius et al. *Pythia8: Color reconnection models*. https://pythia.org/latest-manual/ColourReconnection.html. Accessed: 25.09.2025. 2025.

[11] Spyros Argyropoulos and Torbjörn Sjöstrand. "Color reconnection in $t\bar{t}$ final states". In: *JHEP* 11 (2014). Discusses toy CR models in Pythia including "move", "swap", "flip"; the gluon-move model is one of the considered CR options, p. 043. DOI: 10.1007/JHEP11(2014)043. arXiv: 1407.6653 [hep-ph].

[12] Peter Skands. *RecoilToTop: UserHook in PYTHIA 8 (recoil strategy for top decay)*. CERN, https://skands.physics.monash.edu/slides/pdf/23-Atlas-Pythia.pdf. Described in Pythia documentation and included as UserHook `recoilToTop`; allows the top quark itself

to act as the recoil partner for gluon emissions in the decay. 2023.

[13] The ATLAS Collaboration. "Template-based top mass measurement with alternative recoil models in $t\bar{t}$ ". In: *ATLAS Conf. Note / ATLAS-CONF-2022-058* (2022). Discusses the usage of the recoil-to-top setting in analysis of modelling uncertainties in top mass fits.

[14] Sherpa Team. *Sherpa changelog / release notes (2.2 series)*. https://sherpa-team.gitlab.io/changelog.html. See entries for Sherpa 2.2.1 and 2.2.12 for release-specific notes; 2.2.1 introduced on-the-fly scale variations and no physics changes relative to 2.2.0. accessed 2025.

[15] Johannes Bellm et al. "Herwig 7.1 Release Note". In: *arXiv preprint* (2017). Herwig 7.1 release note; introduces new soft-interaction model, diffractive modelling and a hadronisation retune. arXiv: 1705.06919 [hep-ph].

[16] Johannes Bellm et al. "Herwig 7.2 Release Note". In: *Eur. Phys. J. C* 80 (2020). Herwig 7.2 release note; includes baryonic colour reconnection and hadronisation tune updates, p. 452. DOI: 10.1140/epjc/s10052-020-8011-x. arXiv: 1912.06509 [hep-ph].

[17] Stefan Gieseke et al. "Baryon production from cluster hadronisation". In: *Eur. Phys. J. C* 78 (2018). Describes baryonic colour reconnection / baryon production modelling in Herwig's cluster hadronisation, p. 337. DOI: 10.1140/epjc/s10052-018-5585-7. arXiv: 1803.04961 [hep-ph].

[18] Leif Lönnblad et al. "Baryon correlations in Pythia". In: *arXiv preprint* (2023). Studies baryon pair correlations and baryon-production mechanisms in PYTHIA8; useful to compare baryon modelling between generators. arXiv: 2309.01557 [hep-ph].

[19] ATLAS Collaboration. *Study of baryon energy fraction in jets (ATL-PHYS-PUB-2022-021)*. Tech. rep. ATLAS public note studying baryon energy fraction in jets; provides empirical context for generator baryon-fraction differences. CERN, 2022.

Appendix

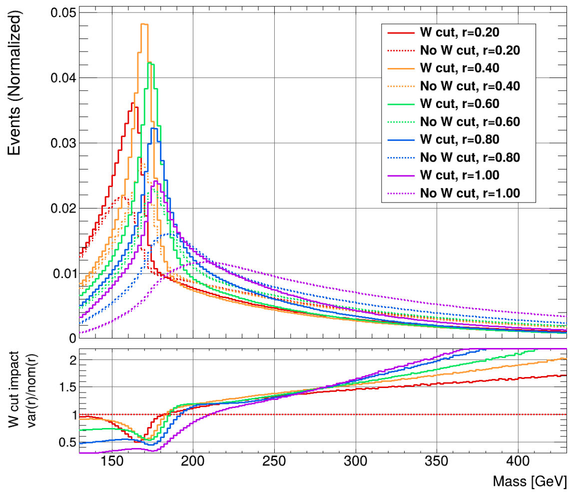

Figure 23: *Top-quark mass with and without the prescribed W boson mass cut.*

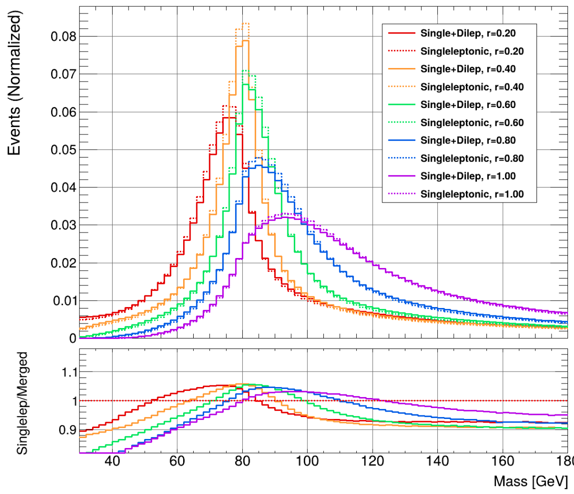

Figure 24: *W boson mass comparison of the singleleptonic* Herwig *7.2.1 dataset and the full dataset.*

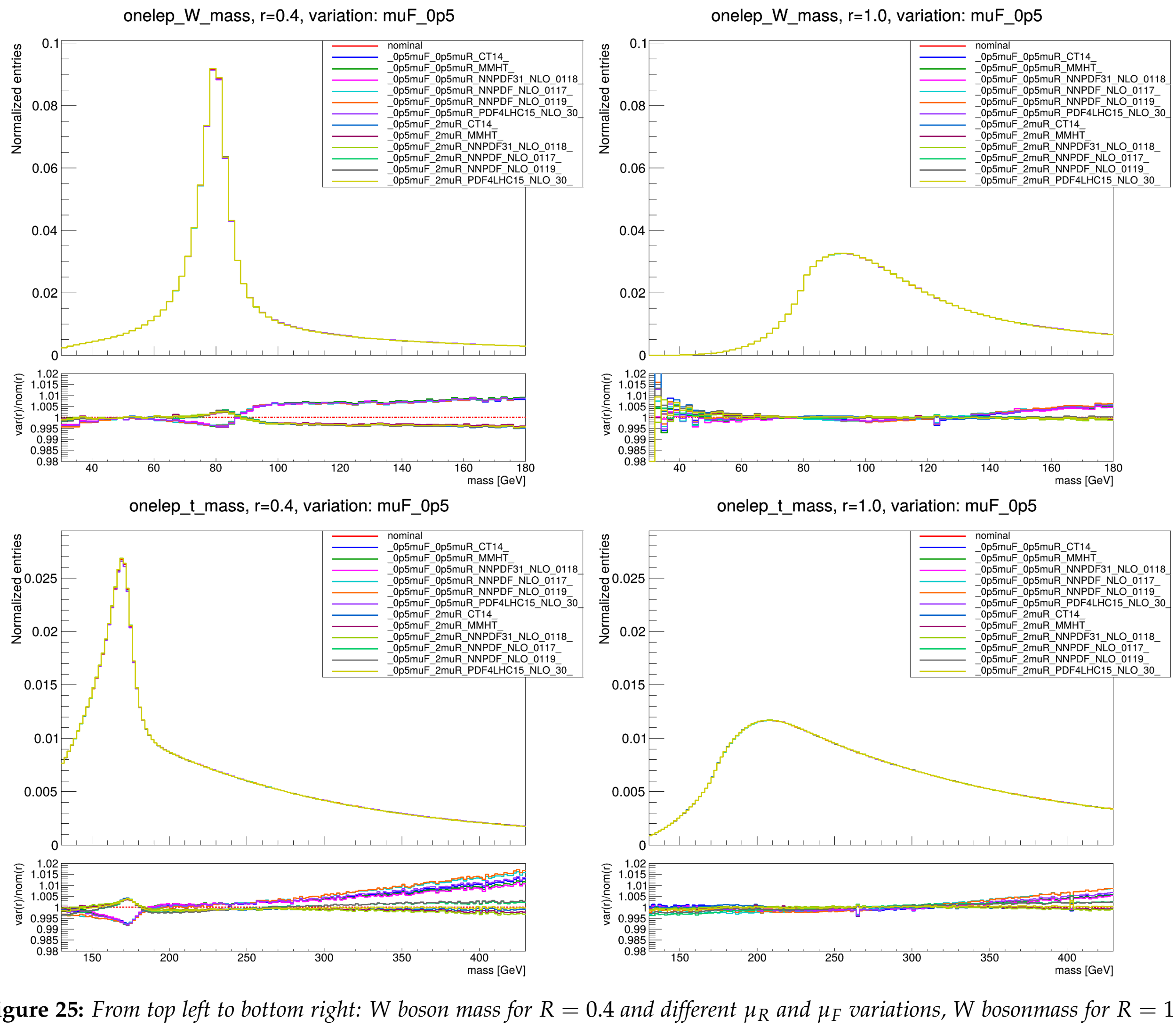

Figure 25: *From top left to bottom right: W boson mass for $R = 0.4$ and different μ_R and μ_F variations, W bosonmass for $R = 1.0$ and different μ_R and μ_F variations, top-quark mass for $R = 0.4$ and different μ_R and μ_F variations, top-quark mass for $R = 1.0$ and different μ_R and μ_F variations.*