ATLAS PUB Note

GROUP-2019-XX

30th September 2019

Draft version 0.1

Study of $t\bar{t}b\bar{b}$ and $t\bar{t}W$ background modelling for $t\bar{t}H$ analyses

The ATLAS Collaboration

This note presents Monte Carlo generator comparisons of the $t\bar{t}b\bar{b}$ and $t\bar{t}W$ processes at particle level. The aim is to compare the modelling of some important backgrounds and the treatment of the associated theory uncertainties for a full Run-2 ATLAS+CMS combination. As a first step, pre-fit modelling and theory uncertainties as used in the experiments are compared in the relevant analysis regions.

Contents

1 Introduction

The search for Higgs boson production in association with a top quark pair ($t\bar{t}H$) in the H $\rightarrow$ bb [1] and H $\rightarrow$ multi-lepton [2] final states is limited by the modelling uncertainties of the main backgrounds, $t\bar{t}b\bar{b}$ and $t\bar{t}W$ respectively. Examples of tree-level diagrams of said processes are shown in Figure 1. A comparison of available Monte Carlo generators is thus performed to study modelling differences. Comparisons of observables are made at particle level, in a phase space similar to the reference measurements. Differences in the object and event selection were made to define a common selection with the CMS collaboration, to later compare these distributions between the two experiments. The goal is to decide on a common strategy between ATLAS and CMS for background modelling uncertainties in the $t\bar{t}H(bb)$ and $t\bar{t}H$(multi-lepton) analyses.
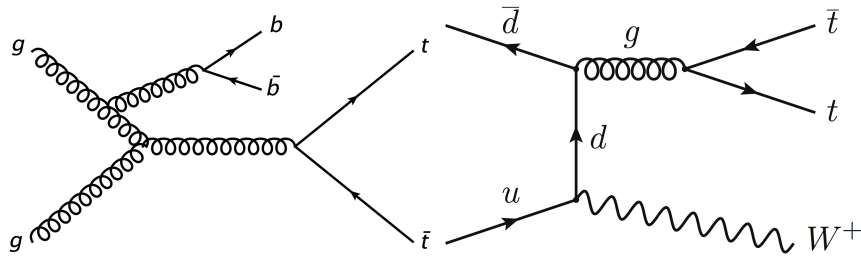

Figure 1: Examples of tree-level Feynman diagrams for $t\bar{t}b\bar{b}$ (left) and $t\bar{t}W$ (right).

2 ttbb

2.1 Samples

Four MC generators are compared in this study, where the inclusive $t\bar{t}$ PP8 sample was previously used as the nominal background estimate in the ttH analysis. It is generated with the POWHEG-BOX v2 NLO event generator [3–6] with NNPDF3.0 NLO PDF set, matched to Pythia8 and is referred to as PP8 $t\bar{t}$, where the additional bb-pair is described by the parton shower. The h_{damp} parameter was set to 1.5 times the top quark mass [7], which is assumed to be 172.5 GeV. The parton shower and the hadronisation were modelled by Pythia 8.210 with the A14 set of tuned parameters. The renormalisation and factorisation scales were set to the transverse mass of the top quark. The intrinsic uncertainty of the nominal PP8 $t\bar{t}$ sample is expressed by the simultaneous variation of the renormalisation and factorisation scales together with the PDF tune parameter. The RadiationUp variation has the renormalisation and factorisation scales decreased by a factor of two, the Var3c upward variation of the A14 parameter set and the h_{damp} parameter doubled. The RadiationDown variation has the renormalisation and factorisation scales increased by a factor of two, the Var3c downward variation and the nominal value of h_{damp}. Additionally, the up and down radiation uncertainty is calculated following the CMS approach , under which the renormalisation scale, factorisation scale and PDF tune variations are each taken individually and their difference to the nominal is summed in quadrature.

The PP8 tt+bb sample also uses the POWHEG generator where $t\bar{t}b\bar{b}$ matrix elements are calculated at NLO with massive b-quarks, using the four-flavour NLO NNPDF3.0 PDF set [8]. The parton shower and hadronisation is modelled by Pythia 8.240.

For the PP8 samples the bottom and charm quark decays are described by EVTGEN v1.2 [9] and the top quark spin correlations follow reference [10].

The SHERPA tt+bb sample describes NLO tt+bb including parton showering and hadronisation by SHERPA-OPENLOOPS [11–13]. The sample was produced with Sherpa version 2.1.1 and the CT10 four-flavour scheme PDF set [14, 15] The renormalisation scale is set to the CMMPS value as in ref [11], the factorisation and resummation scales equal $H_T/2$.

Both the PP8 tt+bb and the SHERPA tt+bb samples describe the additional bb-pair with NLO precision in QCD, taking into account the b-quark mass

The SHERPA $t\bar{t}$ sample uses SHERPA version 2.2.1 [12] with the ME+PS@NLO (multi-leg) setup using the MEPS@NLO prescription [16], interfaced with OPENLOOPS. It provides NLO accuracy for up to one additional parton and LO accuracy for up to four additional partons. The NNPDF3.0 NNLO PDF set is used with a five-flavour scheme and both renormalisation and factorisation scales are set to $\sqrt{0.5 \times (m_{T,t}^2 + m_{T,\bar{t}}^2)}$. A summary of all samples used is given in Table 1.

2.2 Fiducial Volume

Object and event selection is defined at particle-level that closely matches the detector-level described in reference [1]. Jets are reconstructed from stable particles with a mean lifetime of $\tau > 3 \times 10^{-11}$ s, using the anti-k_t algorithm with a radius parameter of $R = 0.4$, and are required to have $p_T > 25$GeV and $|\eta| < 2.5$. Jets are matched to b-hadrons with $p_T > 5$GeV by ghost matching [17] and are referred to as

Table 1: The configurations used for the event generation of ttbb processes.

Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}$	PowHeg v2	NLO	Pythia 8	5FS NNPDF3.0 NNLO	A14
$t\bar{t} + b\bar{b}$	PowHeg v2	NLO	Pythia 8	4FS NNPDF3.0 NNLO	A14
$t\bar{t} + b\bar{b}$	Sherpa 2.1.1	NLO	Sherpa	4FS NNPDF3.0 NNLO	Sherpa default
$t\bar{t}$	Sherpa 2.2.1	tt+0,1pNLO+3p@LO	Sherpa	5FS NNPDF3.0 NNLO	Sherpa default

b-jets. Electrons and muons, referred to as leptons, are required to satisfy $p_T > 27$GeV. Selected leptons are required to be separated from selected jets by $\Delta R > 0.4$. Events are selected with exactly one lepton and at least 4 jets, targeting the semi-leptonic $t\bar{t}$ decay. Two regions are considered, defined by 3 b-jets or $\geq$4 b-jets.

2.3 Results

The nominal PP8 $t\bar{t}$ sample is compared to its radiation uncertainty variations and alternative generators, scaled to a common arbitrary integrated luminosity. The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, where the colour scheme is given in the legend. Discrepancies between PP8 $t\bar{t}$ and the alternative generators can be seen in the ΔR quantities, as in Figures 2 and 3, where at least in the 4b selection the difference to the alternative generators is larger than the uncertainty band given by the radiation variations. In the b-jet pair invariant mass variables $M_{bb}^{maxP_T}$ and $M_{bb}^{min\Delta R}$ shown in Figures 4 and 5, the largest difference is seen between the Sherpa $t\bar{t}$ and all other samples. Differences are also observed in the H_T distributions, particularly in the 3b selection. In H_T of all b-jets, as in Figures 6, one observes a difference between the 4 and 5 flavour schemes, while in H_T of all light jets, shown in Figure 7 a difference between PP8 $t\bar{t}$ and the samples with b-jets in the ME can be seen. The jet multiplicity, as in Figure 8, has poor agreement among the generators for large jet multiplicities. Lastly, differences among the samples are shown for the η_{jet} distribution, in Figure 9.

The second ratio plot shows the relative uncertainty of the radiation variations, shown as the ratio of the uncertainty to the nominal for three cases. The PP8 $t\bar{t}$ sample (black) and the PP8 tt+bb sample (blue), following the above description. It can be seen that the this type of uncertainties is larger for the tt+bb case than in the $t\bar{t}$ calculation. Also, the sum of individual variations following the CMS approach (red) on PP8 $t\bar{t}$ leads to a smaller uncertainty with respect to the simultaneous variation.
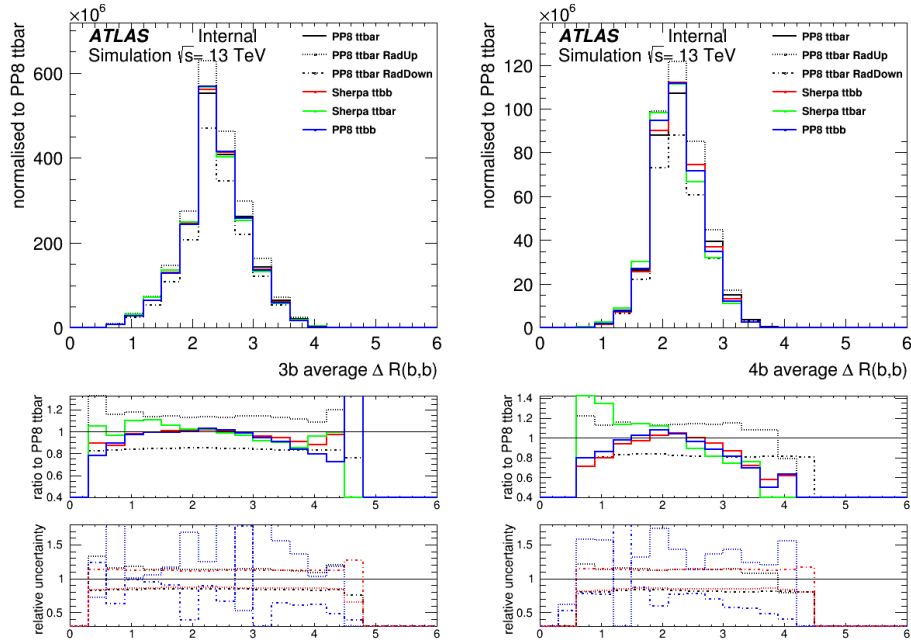
Figure 2: Distribution of the average opening angle between two b-jets, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
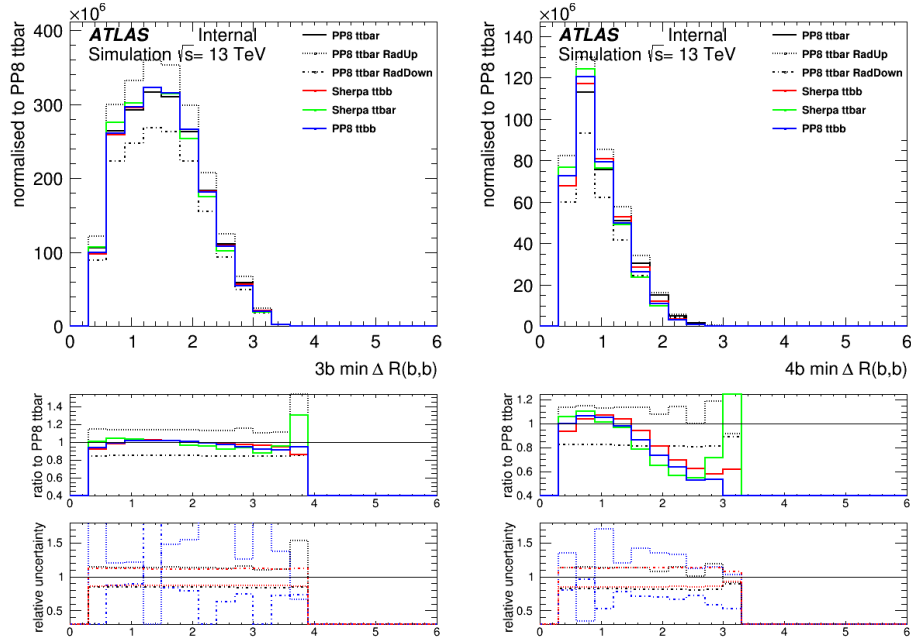

Figure 3: Distribution of the smallest opening angle between two b-jets, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
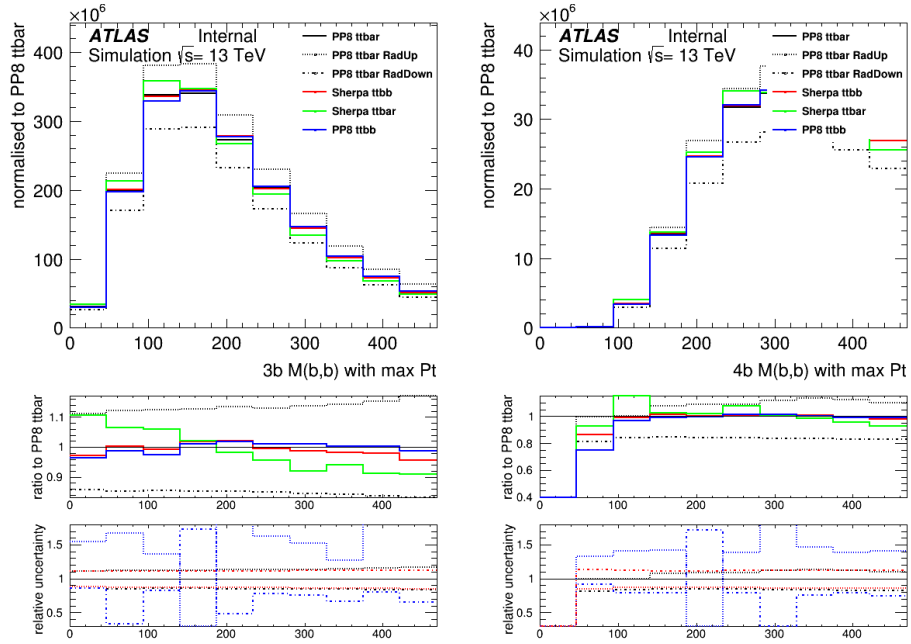
Figure 4: Distribution of the invariant mass in GeV of the two b-jets with the highest P_T sum, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
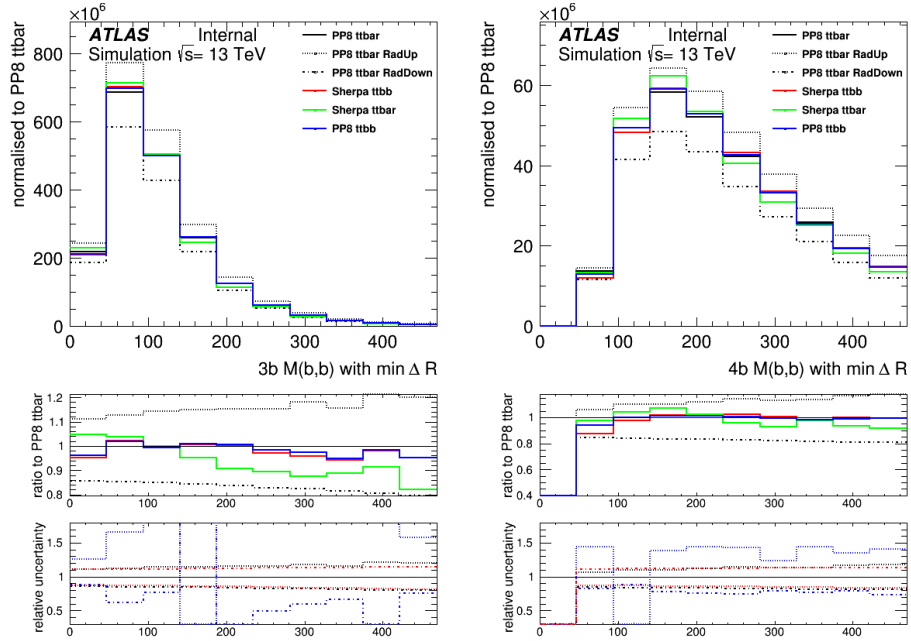

Figure 5: Distribution of the invariant mass in GeV of the two b-jets with the smallest opening angle, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
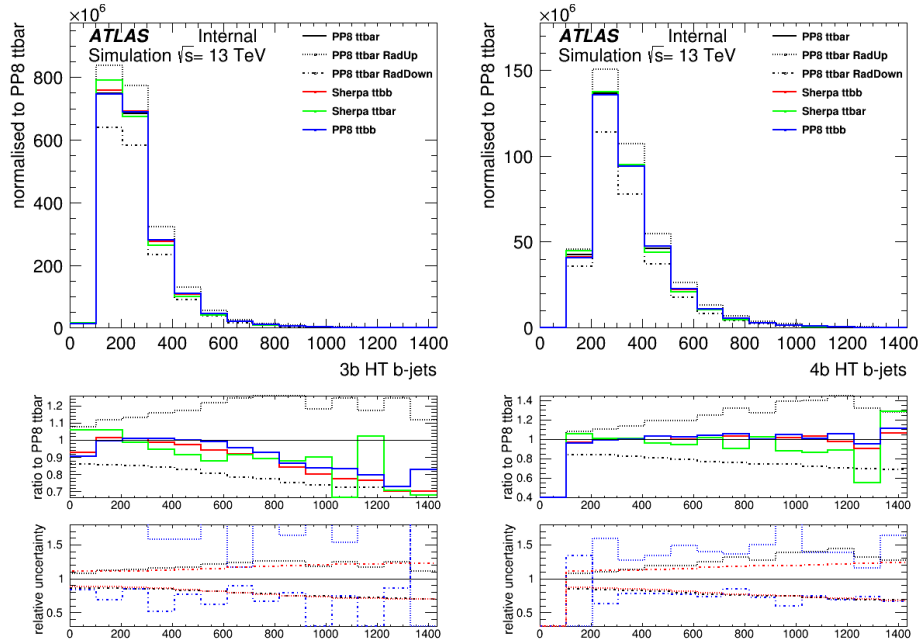
Figure 6: Sum of b-jet transverse momenta in GeV, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
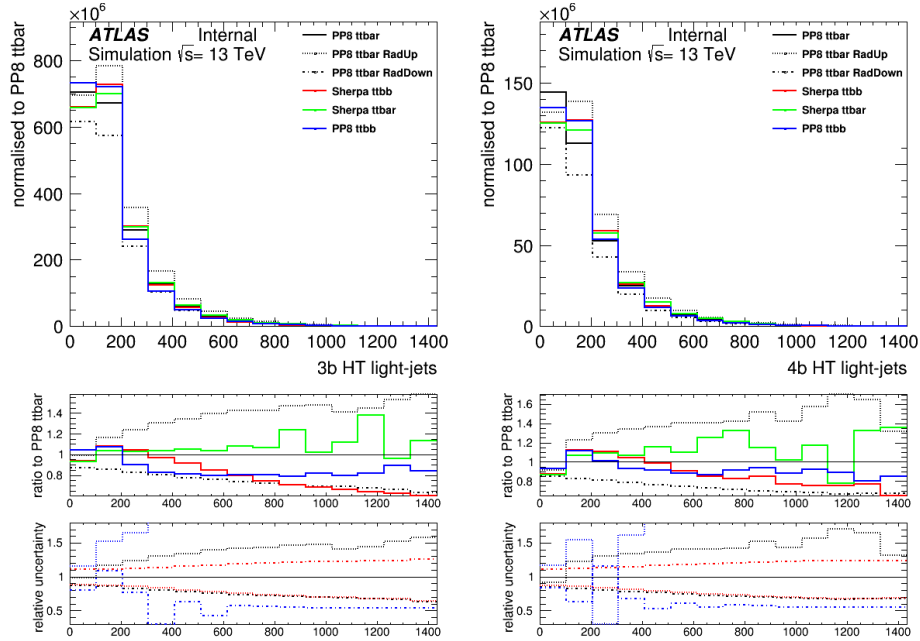

Figure 7: Sum of non-b-jet transverse momenta in GeV, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 $t\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 $t\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 $t\bar{t}$.
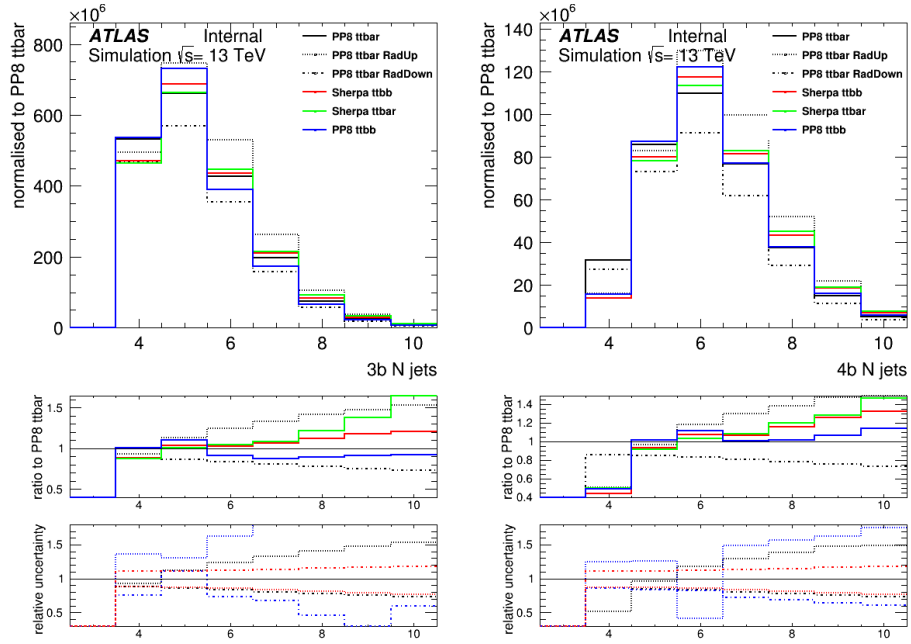
Figure 8: Jet multiplicity, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 t$\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 t$\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 t$\bar{t}$.
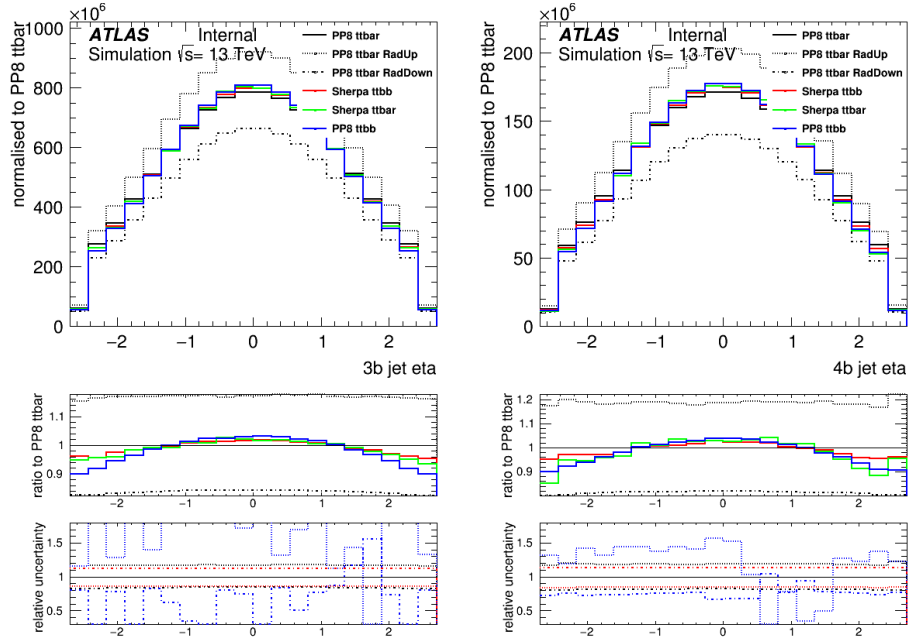

Figure 9: Jet pseudorapidity, for the 3b selection (left) and the 4b-jet selection (right). The first ratio plot shows the ratio of the different MC samples to PP8 t$\bar{t}$, together with its radiation uncertainties. The second ratio plot shows the relative uncertainty of the radiation variations divided by the nominal, for PP8 t$\bar{t}$ (black) PP8 tt+bb (blue), following the above description of simultaneous variations. The sum of individual variations following the CMS approach (red) is shown for PP8 t$\bar{t}$.

Table 2: The configurations used for the event generation of $t\bar{t}W$ processes.

Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}W$	Sherpa 2.2.1	MePs@Nlo	Sherpa	NNPDF3.0 NNLO	Sherpa default
	MG5_aMC	NLO	Pythia 8	NNPDF3.0 NLO	A14

3 $t\bar{t}W$ process

3.1 Samples

Two MC generators are compared in this study. The nominal sample for $t\bar{t}W$ production was generated using the Sherpa 2.2.1 [18] generator with the NNPDF3.0 NLO PDF set. The matrix element (ME) was calculated for up to one additional parton at NLO and up to two partons at LO using Comix [19] and OpenLoops [13], and merged with the Sherpa parton shower [20] using the MePs@Nlo prescription [16]. The choice of renormalisation and factorisation scales is $\mu_R = \mu_F = H_T/2$, where H_T is defined as the scalar sum of the transverse masses $\sqrt{p_T^2 + m^2}$ of all final state particles.

Systematic uncertainties due to missing higher-order QCD corrections are estimated by varying the factorisation and renormalisation scales in the nominal sample simultaneously by a factor of 0.5 and 2.0 with respect to the central value. Uncertainties associated with the modelling of additional QCD radiation are estimated by comparing the nominal $t\bar{t}W$ prediction with that of an alternative sample that was generated at NLO with the MadGraph5_aMC@NLO 2.2.1 generator using the same scale choice and PDF set as for the nominal sample, and interfaced to Pythia 8.2 in combination with the A14 tune. The samples configurations are summarised in Table 2.

3.2 Fiducial Volume

Object and event selection is defined at particle-level that closely matches the detector-level described in reference [2]. Jets are reconstructed from stable particles with a mean lifetime of $\tau > 3 \times 10^{-11}$ s, using the anti-k_t algorithm with a radius parameter of $R = 0.4$. Jets are required to satisfy $p_T > 25$ GeV and $|\eta| < 2.5$. Jets are matched to b-hadrons with $p_T > 5$ GeV by ghost matching [17] and are referred to as b-jets. Electrons and muons, referred to as light leptons, are required to be separated from selected jets by $\Delta R > 0.4$. Hadronically decaying τ leptons are required to satisfy $p_T > 25$ GeV and $|\eta| < 2.5$. Events are selected with exactly two light leptons. Leptons are required to have $|\eta| < 2.5$ and $p_T > 25(20)$ GeV for leading (subleading) lepton. Leptons are required to have same charge, targeting the semi-leptonic $t\bar{t}$ decay and leptonic W decay.

Events with at least 3 jets and least one b-jet are considered in the fiducial volume. The acceptance for events passing this selection is $A_X^{\geq 1b \geq 3j} = 1.79 \times 10^{-2}$ for Sherpa and 1.90×10^{-2} for MadGraph5_aMC@NLO correspondingly. We then split into five regions, categorized by the number of jets (three or ≥ 4), b-jets (one or ≥ 2) as well as the presence of hadronically decaying τ lepton.

Region 1: 1 N_{b-jets}, ≥ 4 N_{jets}, 0-τ_{had}

Region 2: ≥ 2 N_{b-jets}, ≥ 4 N_{jets}, 0-τ_{had}

123   Region 3: 1 N_{b-jets}, 3 N_{jets} , 0-τ_{had}

124   Region 4: ≥2 N_{b-jets}, 3 N_{jets}, 0-τ_{had}

125   Region 5: ≥1 N_{b-jets}, ≥3 N_{jets}, 1-τ_{had}

126   Definitions of the regions are motivated by $t\bar{t}H$ analysis strategy. Regions 1 and 2 corresponds to the
127   signal regions[1] and Regions 3 and 4 are used as control regions in the 2ℓ same-sign 0-τ_{had} $t\bar{t}H$ channel.
128   Definition of Region 5 is closely followed[2] by the selections in 2ℓ same-sign 1-τ_{had} $t\bar{t}H$ channel.

3.3 Results

130   The nominal SHERPA $t\bar{t}W$ sample is compared to its radiation uncertainty variations and the alternative
131   generator. The ratio plot shows the ratio of the alternative MC sample and scale variation to the nominal
132   sample.

133   Sizeable discrepancies in the modelling of jet kinematics can be seen between the SHERPA $t\bar{t}W$ and
134   MADGRAPH5_AMC@NLO generators in 1b-jet selections, while in $\geq 2b$-jets the difference is reduced, as
135   illustrated in Figures 10 and 11 for the high (Regions 1 and 2) and low (Regions 3 and 4) jet multiplicities
136   correspondingly.

137   Good agreement of the single lepton kinematics can be seen between nominal and alternative generators for
138    $\geq 2b$-jets selections, as presented in the right of Figures 12 (presented Region 2, while similar behaviour
139   seen in Region 4 as well). While there is an offset of the order of 10% between nominal and alternative
140   generators for 1b-jet selections, as presented in the left of Figures 12 (Region 1 showed, while similar trend
141   seen in Region 3 as well).

142   Sizeable difference in shapes between nominal and alternative generators observed for the distributions
143   corresponding to the correlations between two leptons such as the angular distance between the two
144   leptons, maximum of lepton's pseudorapidity and azimuthal separation between the leptons, as illustrated
145   in Figures 13 for Region 1 on the right and Region 2 on the left (similar trends observed in Regions 3
146   and 4). As was observed for single lepton kinematics, difference in 1b-jet selections is increased wrt to
147    $\geq 2b$-jets selection.

148   Distributions of the jet multiplicity, number of b-jets, the leading lepton transverse momentum and the
149   angular distance between the two leptons $\Delta R_{\ell\ell}$ for the Region 5 with 1τ_{had} selection are presented on
150   Figure **??**. The difference of 10% observed between nominal and alternative generators for low jet
151   multiplicities, while for $N_{jets} \geq 5$ the difference is covered by the scale variation uncertainties. The
152   distribution of b-jets agrees for N_{b-jets} equal to two, and has sizeable difference in the 1b-jet bin, similarly to
153   0τ_{had} selections. Lepton kinematic distributions has difference in shapes between nominal and alternative
154   generators.

---

[1] slightly different then in [2], in order to define a common selection with the CMS collaboration
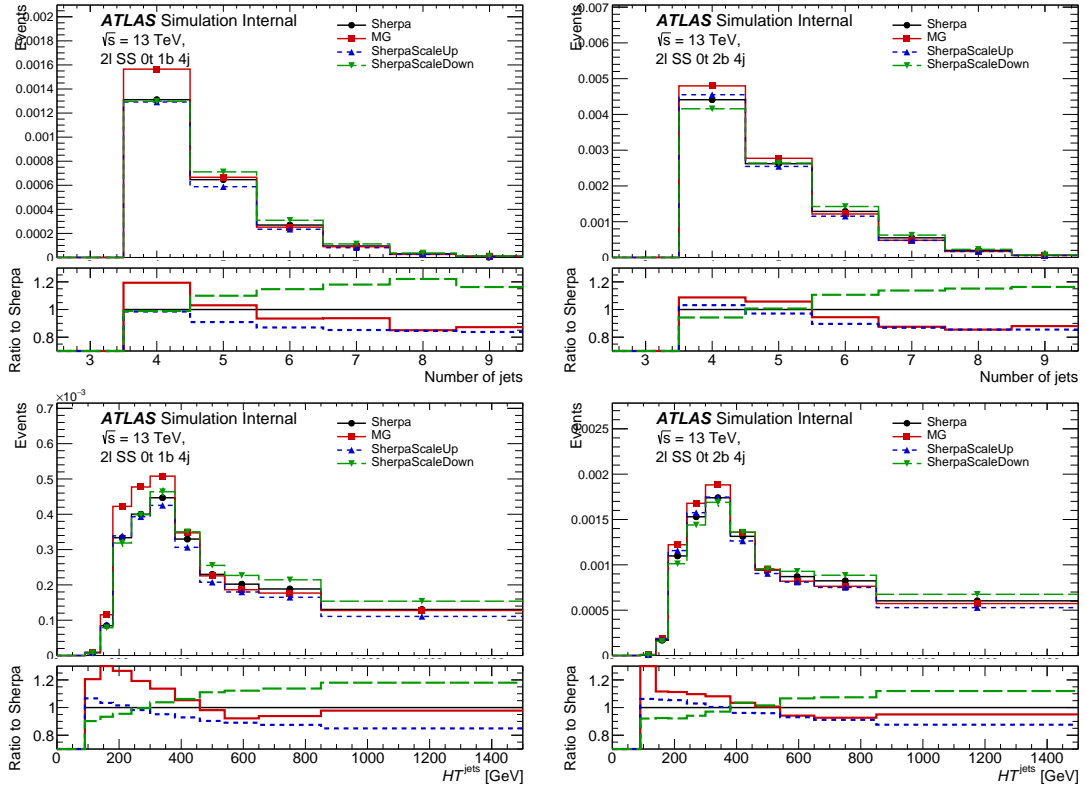[2] requirement on jet multiplicity is relaxed

Figure 10: Distribution of the jet multiplicities (top) and the sum of jets transverse momentum, HT^{jets} (bottom), for the Region 1 with 1b-jet (left) and Region 2 with 2b-jets (right) selection requiring four and more jets. Explanation in text.
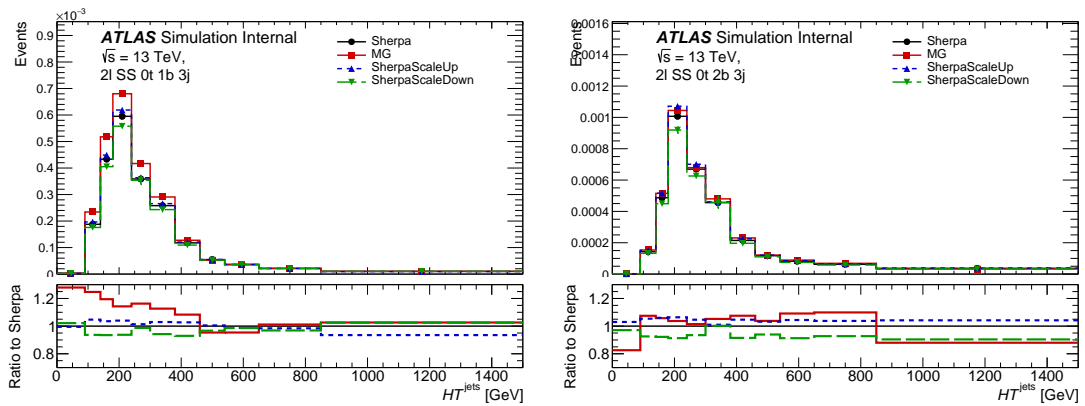

Figure 11: Distribution of the sum of jets transverse momentum, HT^{jets}, for the Region 3 with 1b-jet (left) and Region 4 with 2b-jets (right) selection requiring exactly three jets. Explanation in text.
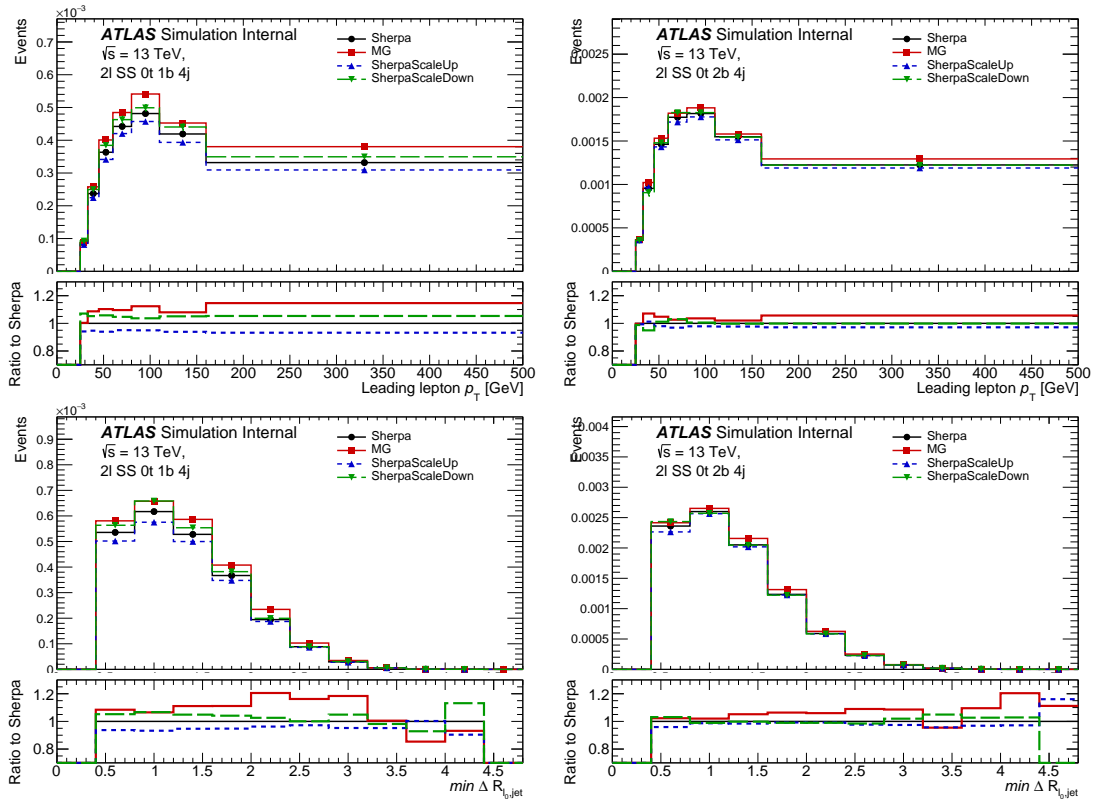
Figure 12: Distribution of the leading lepton transverse momentum (top) and the minimum angular separation between the leading lepton and the nearest jet (bottom), for the Region 1 with 1b-jet (left) and Region 2 with 2b-jets (right) selection requiring four and more jets. Explanation in text.
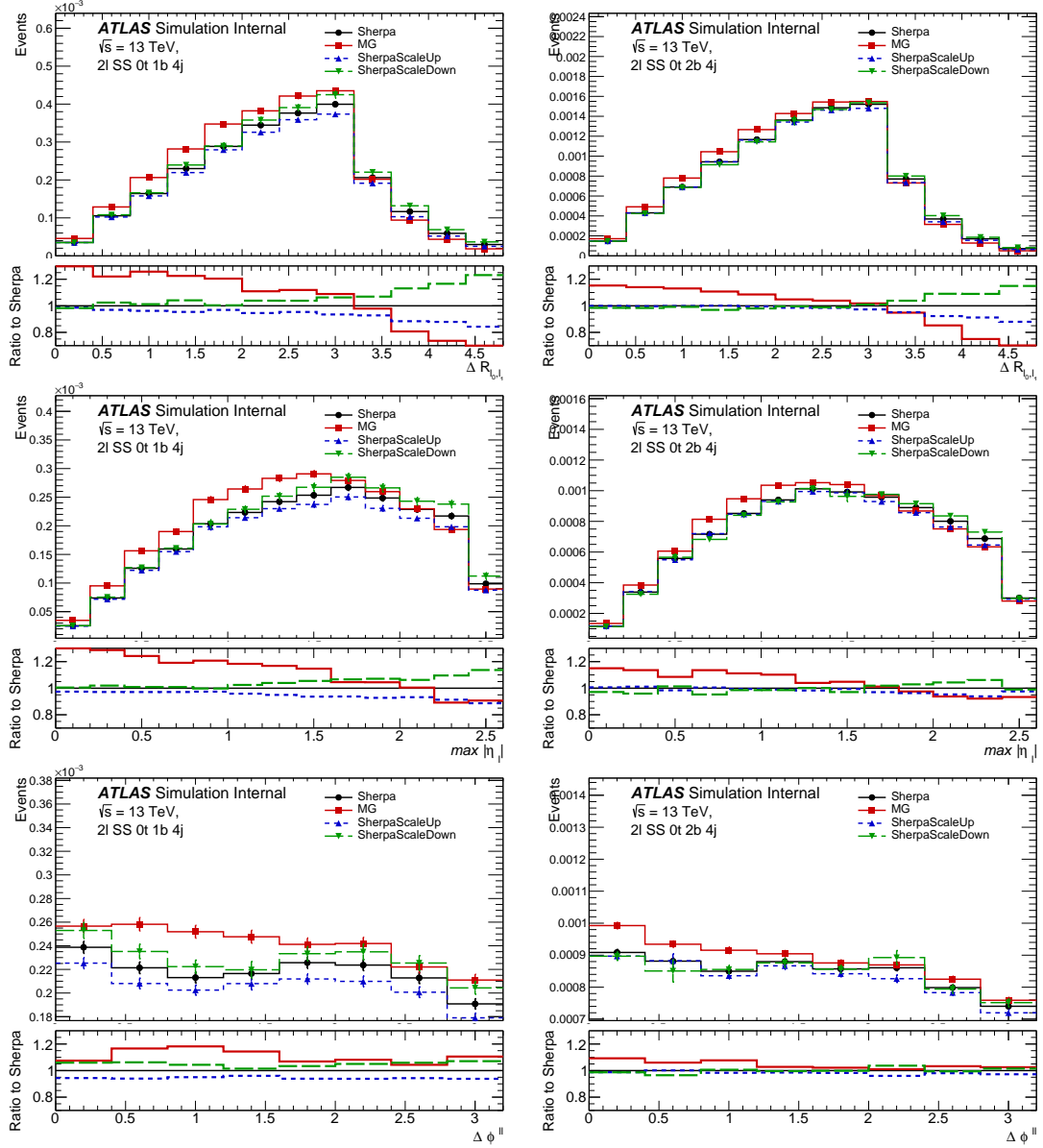
Figure 13: Distribution of the angular distance between the two leptons (top), maximum between lepton $|\eta_{\ell 0}|$ and $|\eta_{\ell 1}|$ (centre), azimuthal separation between the leptons $\Delta\phi_{\ell\ell}$ (bottom) , for the Region 1 with 1b-jet (left) and Region 2 with 2b-jets (right) selection requiring four and more jets. Explanation in text.
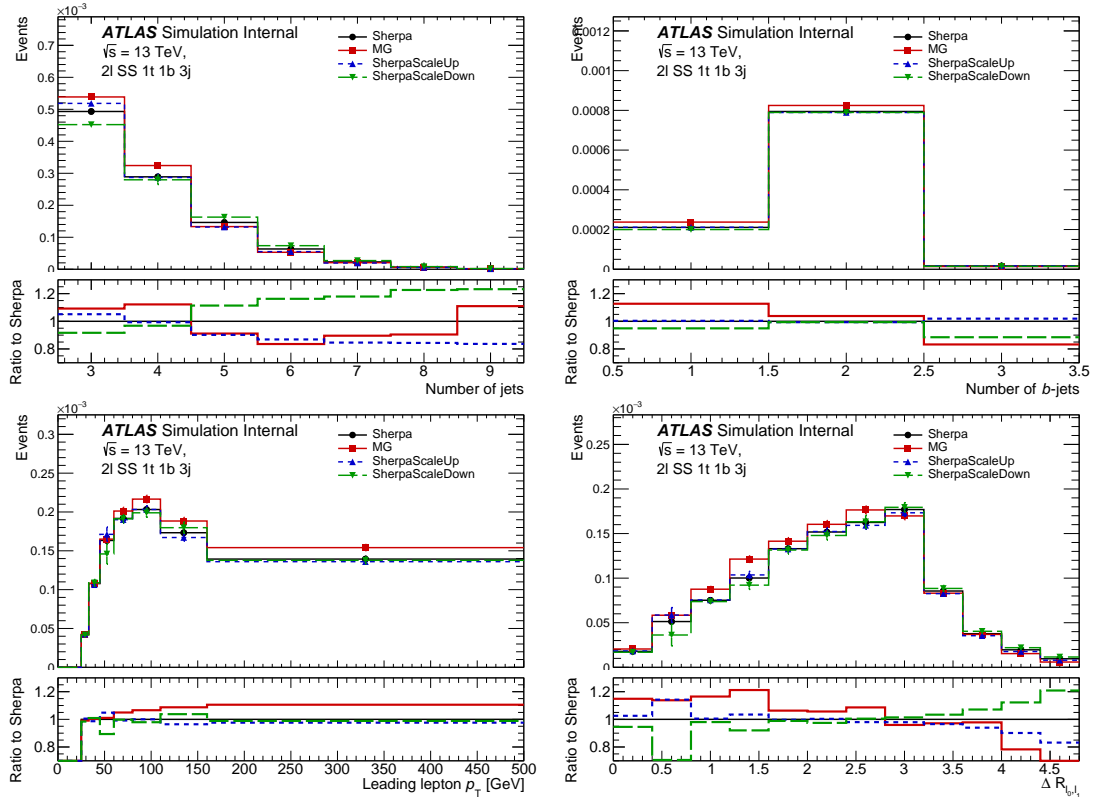
Figure 14: Distribution of the the jet multiplicity, number of b-jets, the leading lepton transverse momentum and the angular distance between the two leptons $\Delta R_{\ell\ell}$ for the Region 5 with $1\tau_{had}$ selection. Explanation in text.

References

[1] ATLAS Collaboration, *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*, Phys. Rev. D **97** (2018) 072016, arXiv: 1712.08895 [hep-ex] (cit. on pp. 3, 4).

[2] ATLAS Collaboration, *Analysis of ttH and ttW production in the multilepton final states with the ATLAS detector*, ATLAS-CONF-2019-045, 2019, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-045/ (cit. on pp. 3, 10).

[3] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 [hep-ph] (cit. on p. 4).

[4] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 4).

[5] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 4).

[6] J. M. Campbell, R. K. Ellis, P. Nason and E. Re, *Top-Pair Production and Decay at NLO Matched with Parton Showers*, JHEP **04** (2015) 114, arXiv: 1412.1828 [hep-ph] (cit. on p. 4).

[7] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168 (cit. on p. 4).

[8] T. Ježo, J. M. Lindert, N. Moretti and S. Pozzorini, *New NLOPS predictions for $t\bar{t}$ + b -jet production at the LHC*, Eur. Phys. J. **C78** (2018) 502, arXiv: 1802.00426 [hep-ph] (cit. on p. 4).

[9] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, ISSN: 0168-9002, URL: http://www.sciencedirect.com/science/article/pii/S0168900201000894 (cit. on p. 4).

[10] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: hep-ph/0702198 [HEP-PH] (cit. on p. 4).

[11] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini and F. Siegert, *NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks*, Phys. Lett. **B734** (2014) 210, arXiv: 1309.5912 [hep-ph] (cit. on p. 4).

[12] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 4).

[13] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on pp. 4, 10).

[14] M. Guzzi et al., *CT10 parton distributions and other developments in the global QCD analysis*, (2011), arXiv: 1101.0561 [hep-ph] (cit. on p. 4).

[15] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. **D89** (2014) 033009, arXiv: 1302.6246 [hep-ph] (cit. on p. 4).

[16] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on pp. 4, 10).

[17] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph] (cit. on pp. 4, 10).

[18] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 10).

[19] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 10).

[20] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 10).