Simulation of top-quark production
for the ATLAS experiment at $\sqrt{s} = 13$ TeV

The ATLAS Collaboration

Abstract

This note summarizes the Monte Carlo simulation setup for the pair and single production
of top quarks for the ATLAS experiment at the LHC for $\sqrt{s} = 13$ TeV. In addition to the
settings available and recommended for analyses using the 2015 dataset, the anticipated
setup for 2016 data analysis is also discussed.
1 Introduction

The modeling of final state particles is one of the major sources of systematic uncertainties in measurements of top quark cross sections and of top quark properties in proton-proton collisions at the LHC. Since top quark pair (t\bar{t}) production is one of the major backgrounds for Higgs measurements and searches for physics beyond the Standard Model, these analyses would also benefit from a better understanding of t\bar{t} modeling.

This note is organized as follows. Section 2 gives a short description of the measurements used in this note for making comparisons between data and simulation. The comparisons with data are an update of the t\bar{t} modeling studies published in 2015 based on parameter variations within the Powheg+Pythia6 setup [1] and a more general set of generator comparisons [2] and use partly the same data measurements. The comparisons are performed within the Rivet framework [3] v2.2.0. The nominal simulation setup and the samples for systematic uncertainties used by analyses on the 2015 dataset are discussed in Sec. 3. Studies to prepare the simulation for analysis in 2016 are documented in Sec. 4. Additional considerations and current findings for the various simulation setups are discussed in more detail in Sec. 5.

2 Data measurements and Rivet routines

In this note predictions of various Monte Carlo (MC) generators are compared to data. Because the number of unfolded distributions currently available with Run II data is limited, determination of optimal settings of MC generator parameters is performed using measurements from Run I. It is important to control the real emission from the NLO calculation and further emissions from the parton shower. The hardest emission will recoil against the t\bar{t} system and is expected to influence the transverse momenta of the t\bar{t} system and jets. These quantities will be used as benchmark distributions in the generator comparisons to find an appropriate systematic coverage. In addition the impact of parameter and generator variations sensitive to the initial and final state radiation on the top-quark p_T and the jet multiplicity are studied. The effect of the choice for the parton density functions is studied for the rapidity of the t\bar{t} system. The analyses that are used to reconstruct these distributions are presented in the following.

- Analysis A: Measurement of the t\bar{t} production cross-section as a function of jet multiplicity and jet transverse momentum at \( \sqrt{s} = 7 \) TeV (Rivet analysis: ATLAS_2014_I1304688 [4])

  The differential cross-sections are measured as a function of the jet multiplicity for up to eight jets using jet transverse momentum thresholds of 25, 40, 60, and 80 GeV, and as a function of jet transverse momentum up to the fifth leading jet [4]. The events are selected in the lepton+jets channel, requiring exactly one charged lepton and at least three jets with at least one jet which is identified to originate from a b-quark. Jets are reconstructed using the anti-\( k_T \) jet algorithm with a radius parameter \( R = 0.4 \) [5] and are required to have a \( p_T > 25 \) GeV and \(|\eta| < 2.5\). Jets originating from b-quarks are defined using a B-hadron matching, requiring at least one B-hadron to be found within the jet using the ghost association technique [6, 7, 8]. The results are defined at truth particle level and are corrected for all detector effects, within a kinematic range closely matched to the experimental acceptance. The data used in this measurement were taken at the ATLAS experiment at \( \sqrt{s} = 7 \) TeV.
• **Analysis B: Measurement of normalized differential cross-sections for $t\bar{t}$ production using the reconstruction of pseudo-top quarks at 7 TeV (Rivet analysis: ATLAS_2015_I1345452 [9])**

Differential cross-sections are measured in $t\bar{t}$ events at a center-of-mass energy of $\sqrt{s} = 7$ TeV [9] in events with exactly one charged lepton and at least four jets with $p_T > 25$ GeV. At least two of the jets have to be identified as $b$-jets. The top quarks are reconstructed from truth level leptons, jets and neutrinos, using model-independent pseudo-top quark definitions to reconstruct the hadronic and leptonic top quarks. The measurements are performed as function of the $p_T$ and $\eta$ of the top quarks and the $t\bar{t}$ system.

• **Analysis C: Single top reconstruction for the $t$-channel production in single lepton events**

This Rivet analysis is not based on collision data and is used to compare generator predictions at $\sqrt{s} = 13$ TeV. Charged leptons, neutrinos and jets with a radius parameter of 0.4 are selected.

The $W$ boson is reconstructed from the lepton-neutrino combination which has an invariant mass closest to the $W$ mass of 80.4 GeV. The $W$-boson candidate is then used together with the selected jets to reconstruct the top quark candidate. Only jets containing $B$-hadrons are used here. The jet-$W$ combination which has the closest invariant mass to the top-quark mass of 172.5 GeV is considered as the top-quark candidate.

3 Nominal and systematic samples

In the following, nominal and systematic samples for $t\bar{t}$ and single-top production are described. The nominal samples for $t\bar{t}$ and single-top production are produced with Powheg+Pythia6 and compared with different matrix-element plus parton-shower (ME+PS) setups described in Tab. 1. The distributions are compared to 7 TeV ATLAS data where the corresponding samples and unfolded data are available. In addition, direct MC-to-MC comparisons are performed using MC samples produced at a center-of-mass energy of 13 TeV. The modeling of the signal processes depends on many aspects (ME generation, parton shower, hadronization) and in general when performing an analysis that relies on $t\bar{t}$ MC, uncertainties on the different aspects are assessed in an independent way. Table 1 depicts a possible breakdown of modeling uncertainties together with the samples that can be used to evaluate each uncertainty component. Further details about the generator setups are given in Sec. 3.1.

3.1 Simulation setup for analyses with the 2015 dataset

A summary of the generator settings for the $t\bar{t}$ MC samples discussed in this Section are presented in Tab. 2. If not stated otherwise, the same settings are used for the generation of the single top samples shown in this document.

• The Powheg+Pythia6 sample is generated using the Powheg NLO generator [10, 11, 12] (revision 2330, version 3.0) with the NLO CT10 [13] PDF. A NLO matrix element is used for the $t\bar{t}$ hard scattering process. The resummation damping factor $h_{\text{damp}}$, which is one of the parameters controlling the ME/PS matching in Powheg and effectively regulates the high-$p_T$ radiation, is set to the top-quark mass. The PS, hadronization and the underlying event (UE) are simulated using Pythia6 [14] (version 6.427) with the Perugia 2012 tune [15] and the corresponding leading order (LO) CTEQ6L1 PDFs [16].
The hard process renormalization $\mu_r$ and factorization $\mu_f$ scales are set to the generator default value $\mu$ that is defined by:

$$\mu = \sqrt{m_{\text{top}}^2 + p_T^2}. \quad (1)$$

where $m_{\text{top}}$ and $p_T$ are the top-quark mass and the top-quark transverse momentum, evaluated for the underlying Born configuration (i.e. before radiation).

Additional samples are produced to estimate the effect of more or less radiation, with variations on the factorization and renormalization scale, the used $h_{\text{damp}}$ value and the Perugia 2012 radiation tune variations.

- The Powheg+Pythia8 sample is generated using the same settings for Powheg as for the Powheg+Pythia6 sample but the PS, hadronization and the UE are simulated using Pythia8 [17] (version 8.183) with the A14 tune [18] and the corresponding CTEQ6L1 PDFs. The matching of the ME to the PS is performed using the Main31 user hook [19] of Pythia8 using the settings listed in Tab. 3. The parameter $n_{\text{Final}}$ sets the number of particles that are produced at Born level before any radiation and is set to 2. The parameter $p_{\text{Thard}}$ is set to the factorization scale ($p_{\text{Thard}}=0$).

- The Powheg+Herwig++ sample is generated using the same setup for Powheg as for the Powheg+Pythia6 sample but the PS, hadronization and the UE are simulated using Herwig++ [20] (version 2.7.1) with the UE-EE-5 tune [18] and the corresponding CTEQ6L1 PDFs.

- The MadGraph5_aMC@NLO+Herwig++ sample is generated with version 2.2.1 of the MadGraph5_aMC@NLO event generator [21]. A NLO ME and CT10 PDF are used for the $t\bar{t}$ hard scattering process. The renormalization and factorization scales are set to the transverse mass $\mu = \sqrt{m_T^2 + 0.5 \cdot (p_{T,t}^2 + p_{T,\bar{t}}^2)}$. The PS, hadronization and the UE are modeled using the Herwig++ (version 2.7.1) [20] generator. The UE-EE-5 author Herwig++ tune and the corresponding CTEQ6L1 PDF are used.

- The MadGraph5_aMC@NLO+Pythia8, sample is generated with version 2.2.1 of the MadGraph5_aMC@NLO event generator [21]. A NLO ME and NNPDF3.0 PDF are used for the $t\bar{t}$ hard scattering process. The renormalization and factorization scales are set to $H_T/2$. The PS, hadronization and the UE are modeled using the Pythia8 [17] (version 8.183) generator. The A14 tune and the corresponding NNPDF2.3LO PDF are used.

- The Sherpa [22] sample is generated using v2.1.1 with the MEPS@NLO [23] setup. The events are generated with a $t\bar{t}$ matrix element plus 0 and 1 parton simulated at NLO and 2, 3 and 4 partons at LO. The CT10 PDF is used and the PS, hadronization and UE are simulated using the default Sherpa settings. For this sample, no tuning to ATLAS data has been performed. This generator version is known to not describe the jet quantities very well. A newer version of the generator will allow to get a better agreement in the future.

All MC generator samples are produced with $m_{\text{top}} = 172.5$ GeV and within the ATLAS software framework [24].

\[ ^1 H_T \text{ is defined as the scalar sum of the transverse masses using all final state particles. } \]
Table 1: A summary of the modeling systematic uncertainties for the $t\bar{t}$ and single top processes as well as the samples used in each case. The symbol $\Delta$ denotes the difference in the analysis observables using the simulation from the samples column. The notation $\pm|\Delta|$ indicates that the full difference is symmetrized and applied to the nominal. Without the absolute value bars, $\Delta$ indicates that the signed difference with respect to the nominal for all variations is used to estimate the uncertainty. Details about the generator setups used are given in Sec. 3.1.

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>Samples</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>POWHEG+PYTHIA6</td>
<td>N/A</td>
</tr>
<tr>
<td>Hard Scatter Generation</td>
<td>POWHEG+HERWIG++ vs. aMC@NLO+HERWIG++</td>
<td>$\pm</td>
</tr>
<tr>
<td>Parton Shower and Hadronization Model</td>
<td>POWHEG+PYTHIA6 vs. POWHEG+HERWIG++</td>
<td>$\pm</td>
</tr>
<tr>
<td>Scales and Additional Radiation</td>
<td>POWHEG+PYTHIA6 Variations</td>
<td>$\Delta$</td>
</tr>
<tr>
<td>Interference ($Wt$ and $t\bar{t}$)</td>
<td>POWHEG+PYTHIA6 DR vs. DS Schemes</td>
<td>$\pm</td>
</tr>
</tbody>
</table>

Table 2: ME and PS/UE generator settings for each of the MC samples used for the 2015 analyses. The generator versions and the PDFs used for the ME and in the PS are shown alongside the tune.

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>ME Gen.</th>
<th>PS/UE Gen.</th>
<th>ME &amp; PS/UE PDF</th>
<th>PS Tune</th>
<th>Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>POWHEG+PYTHIA6</td>
<td>POWHEG-Box r2330.3</td>
<td>PYTHIA 6.427</td>
<td>CT10 &amp; CTEQ6L1</td>
<td>P2012</td>
<td>Powheg Matching</td>
</tr>
<tr>
<td>POWHEG+PYTHIA8</td>
<td>POWHEG-Box r2330.3</td>
<td>PYTHIA 8.183</td>
<td>CT10 &amp; CTEQ6L1</td>
<td>A14</td>
<td>Main31 (pThard=0 nFinal=2)</td>
</tr>
<tr>
<td>POWHEG+HERWIG++</td>
<td>POWHEG-Box r2330.3</td>
<td>HERWIG++ 2.7.1</td>
<td>CT10 &amp; CTEQ6L1</td>
<td>UE-EE-5</td>
<td>Powheg Matching</td>
</tr>
<tr>
<td>aMC@NLO+HERWIG++</td>
<td>aMC@NLO 2.2.1</td>
<td>HERWIG++ 2.7.1</td>
<td>CT10 &amp; CTEQ6L1</td>
<td>UE-EE-5</td>
<td>MC@NLO</td>
</tr>
<tr>
<td>aMC@NLO+PYTHIA8</td>
<td>aMC@NLO 2.2.1</td>
<td>PYTHIA 8.183</td>
<td>NNPDF3.0 NNPDF2.3LO</td>
<td>A14</td>
<td>MC@NLO</td>
</tr>
<tr>
<td>SHERPA</td>
<td>SHERPA 2.1.1</td>
<td>SHERPA</td>
<td>CT10</td>
<td>Default</td>
<td>MEPS@NLO (Q=30 GeV)</td>
</tr>
</tbody>
</table>
3.2 Radiation variations with Powheg+Pythia6

The impact of hard scatter scale variations and additional radiation from the variation of the $h_{\text{damp}}$ parameter are studied in the following using the Powheg+Pythia6 generator setup as described above. The nominal sample is compared with two variations:

- The factorization and hadronization scales are varied by a factor of 0.5, and, simultaneously, the $h_{\text{damp}}$ parameter is increased to $2m_{\text{top}}$. Furthermore the radHi tune variation from the P2012 tune is used. For the single-top $t$-channel samples, the $h_{\text{damp}}$ parameter is not changed.

- The factorization and hadronization scales are varied by a factor of 2.0, while the $h_{\text{damp}}$ parameter is unchanged. Furthermore the radLo tune variation from the P2012 tune is used.

These samples have been studied before in Ref. [1] using different unfolded distributions and have found to envelope the setups where $\mu_R$ and $\mu_F$ are varied independently. Their performance is now revised based on newer measurements that allow a full reconstruction of the $t\bar{t}$ final state. The distributions are compared to data for $t\bar{t}$ events using different observables as shown in Fig. 1 and MC-to-MC comparisons are made for the single top $t$-channel production in the lepton+jets channel as shown in Fig. 2.

For the comparison to data, the number of jets with a $p_T$ larger than 25 GeV and the $p_T$ of the leading jet are evaluated in Fig. 1(a) and (b), respectively. Furthermore the transverse momentum of the reconstructed hadronic top and the $t\bar{t}$ system are studied in Fig. 1(c) and (d). The smallest impact is observed for the $p_T$ of the hadronic top candidate and the leading jet $p_T$, while the jet multiplicity and the $p_T$ of the $t\bar{t}$ spectrum are very sensitive to these variations and bracket the experimental uncertainties.

For the $t$-channel samples, the same parameter variations are used. The corresponding distributions are shown in Fig. 2. The jet and top-quark momentum have very low sensitivity to the variations while the jet multiplicity shows symmetric changes of up to 6%.

The generator setups used here to study radiation systematics have been shown in Ref. [1] to provide symmetric up and down variations that cover the measured uncertainties on the data for Analysis A and the jet gap observables [25] in dilepton $t\bar{t}$ events. In this document the studies are repeated while using newer unfolded data distributions in the lepton+jets channel (Analysis B). Similar conclusions can be drawn and hence these samples are used to estimate the radiation systematic for the 13 TeV analyses for $t\bar{t}$ and single top.
Figure 1: The Powheg+Pythia6 samples with different factorization and renormalization scale variations and different $h_{\text{damp}}$ values are compared to the nominal sample and data at $\sqrt{s} = 7$ TeV. The comparison is performed for the jet multiplicity (a) and the leading jet momentum (b) from Analysis A [4] as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level in Analysis B [9]. The data are represented as closed (black) circles with statistical uncertainties. The yellow band is the total experimental uncertainty on the data (statistical and systematic). The generator predictions are shown as solid and dashed colored lines.
Figure 2: The Powheg+Pythia6 samples with different factorization and renormalization scale variations and different $h_{damp}$ values are compared to the nominal sample at $\sqrt{s} = 13$ TeV. The comparison is performed using Analysis C for the jet multiplicity (a) and the leading jet momentum (b) as well as for the reconstructed leptonic top-quark $p_T$ (c) in single top $t$-channel events in the lepton+jets channel.
3.3 Generator comparisons

In the following, $t\bar{t}$ samples produced with different ME generator and different PS setups are compared to the default Powheg+Pythia6 sample to estimate generator modeling uncertainties and evaluate the performance of alternative generator setups. The effect of parton showering and hadronization is estimated by comparing the Powheg+Pythia6 sample with Powheg+Herwig++. Large deviations are observed in Fig. 3 for the jet multiplicity distribution and the jet and top-quark transverse momenta. The deviations are very prominent for the low $p_T$ values, where the Powheg+Herwig++ sample shows a large disagreement with the unfolded data. This effect is further discussed in Sec. 5.1. The ME only controls the radiation up to the fifth jet in lepton+jets events. Studies of higher jet multiplicity therefore are sensitive to the choice of parameters in the shower generation. In addition, the samples are compared to Powheg+Pythia8. The data/MC agreement for the jet multiplicity and the transverse momentum of the $t\bar{t}$ system is not as good as for Powheg+Pythia6 and will be studied in more detail in Sec. 4.
Figure 3: Comparison of different generator setups produced at $\sqrt{s} = 7$ TeV for the jet multiplicity (a) and the leading jet momentum (b) (Analysis A) as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events (Analysis B). The data and generator predictions are represented in the same way as in Fig. 1.
The uncertainty on the hard scattering generation is evaluated by comparing the Powheg+Herwig++ and the aMC@NLO+Herwig++ sample as shown in Fig. 4. The former predicts a higher number of jets than the alternative sample and the data. It furthermore predicts softer transverse momentum distributions, leading to differences between the two samples of up to 20% in the benchmark distributions. When comparing an alternative setup using Powheg+Pythia8 and aMC@NLO+Pythia8, the differences are smaller for the momenta of the hadronic top quark and the $t\bar{t}$ system.

![Graphs](image)

Figure 4: Comparison of different generator setups produced at $\sqrt{s} = 13$ TeV for the jet multiplicity (a) and the leading jet momentum (b) (Analysis A) as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events (Analysis B).
3.4 EvtGen studies

For all simulated top-quark processes in Run II that use PYTHIA or HERWIG for the parton shower and hadronization, EvtGen 1.2 [26] was used in the production. After PYTHIA or HERWIG is run in the simulation chain, the decay products from heavy-flavor hadrons are removed from the event record and these hadrons are then re-decayed using EvtGen. The impact of EvtGen on low-level quantities within jets was studied extensively in Ref. [27]. Figure 5 shows the impact of EvtGen on the b-tagging efficiency at $\sqrt{s} = 8$ TeV, which is important for many analyses involving top quarks. The efficiency to tag a jet originating from a b-quark as such depends strongly on the number, type and momentum of particles produced in the decay of heavy-flavor hadrons. Therefore generator dependent b-tagging scale factors have been used in this case. These quantities are harmonized with EvtGen and so the spread in the b-tagging efficiency for the various generators is significantly reduced.

![Graphs showing b-tagging efficiency](image)

Figure 5: The efficiency to tag jets originating from b-quarks as b-jets using a requirement on the MV1 multivariate discriminant [28] that is chosen such that it is 70% efficient in $t\bar{t}$ events without EvtGen (a) and with EvtGen (b). For both plots, the jet $p_T$ and $\eta$ distributions are re-weighted to remove the kinematic-induced differences in the tagging efficiency between the generators. These simulations are performed at 8 TeV but use the default Run 2 13 TeV settings described in Sec. 3.
4 Studies of Powheg+Pythia8

In this section, studies using Powheg+Pythia8 $t\bar{t}$ samples are discussed. The data/MC agreement of the Powheg+Pythia8 sample introduced in Sec. 3.1 is studied to find an optimal configuration for the nominal sample. The samples are produced with the A14 tune [18] and the impact of the corresponding eigentunes are evaluated by comparing the simulated samples to unfolded collision data using the Rivet framework [3]. The studies are performed using $t\bar{t}$ events and the measurements of Analysis A and B introduced in Sec. 2.

4.1 Nominal sample for Powheg+Pythia8

For the Powheg+Pythia8 sample discussed in Sec. 3, the data/MC agreement is worse than for the Powheg+Pythia6 sample used as a default. To improve the description of the data, the parameter settings for the Main31 user hook have been varied and the setup with the best performance has been chosen as new default for Powheg+Pythia8 that will be used in future analyses. Furthermore, NNPDF3.0 has been chosen for the PDF used in the ME calculation to improve the description of the $t\bar{t}$ rapidity distribution as shown in Ref. [2]. The old and new Main31 parameters are shown in Tab. 3. The parameter nFinal, which sets the number of particles that are produced at Born level before any radiation, is set to 2. The parameter pThard is set to the factorization scale shown in Eq. 1. These settings were used in both samples compared here. To improve the data modeling, the pTdef parameter that describes the definition of the transverse momentum was changed from the Powheg to the Pythia $p_T$ definitions. Furthermore the transverse momentum of the emission was set to the $p_T$ of the emitted parton with respect to the radiating parton (pTemt parameter). Comparisons of kinematic distributions obtained using the two generator setups are shown in Fig. 6. The benchmark distributions show improved agreement between MC and data with the new settings, especially for the jet multiplicity and top-quark $p_T$ and the new settings will be used for future Powheg+Pythia8 $t\bar{t}$ samples. The change of the PDF choice in the ME calculation results in the expected improved description of the $t\bar{t}$ rapidity distribution, as shown in Fig. 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Old Setup</th>
<th>New Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTdef</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>pTemt</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PDF</td>
<td>CT10</td>
<td>NNPDF3.0</td>
</tr>
</tbody>
</table>

Table 3: Parameter settings for Main31 in the old Powheg+Pythia8 sample (left) and the updated settings (right).

The systematic variations for the A14 tune$^2$ are evaluated using the provided eigentune variations. Five up and down variations are provided to estimate the impact of multi-parton interaction (MPI), initial-state radiation (ISR) and final-state radiation (FSR). The effects of these variations are discussed in Sec. 4.2 and 4.3 respectively.

$^2$Parameters varied for each eigentune are listed in Ref. [18] Tab. 4.
$$d\sigma/dn_{\text{jets}} \ [pb]$$

<table>
<thead>
<tr>
<th>njets (jet $p_T &gt; 25$ GeV)</th>
<th>Expected</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.8</td>
<td></td>
</tr>
</tbody>
</table>

$$d\sigma/dp_T \ [pb/GeV]$$

<table>
<thead>
<tr>
<th>$p_T (\hat{t}h)$ (GeV)</th>
<th>Expected</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6: The default Powheg+Pythia8 sample is compared with a Powheg+Pythia8 sample where the settings in the Main31 user hook are varied. The comparison is performed using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV for the jet multiplicity (a) and the leading jet momentum (b) (Analysis A [4]) as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
Figure 7: The default Powheg+Pythia8 sample is compared with a Powheg+Pythia8 sample where the settings used in the Main31 user hook are varied. The comparison is performed for the absolute rapidity of the $t\bar{t}$ system in $t\bar{t}$ $e$+jets events (a) and $\mu$+jets events (b) using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
4.2 Multiple parton interaction and color reconnection

The systematic variations for the A14 tune [18] are evaluated in the following. The samples containing the MPI parameter variations (Var1) are compared with the nominal sample and the unfolded data in Fig. 8. The impact on the jet multiplicity is largest for a large number of jets (up to 10%) but negligible for the momentum distributions of the jet and the top/tt-system.

Figure 8: The impact of multiparton interaction and color reconnection variations corresponding to the Powheg+Pythia8 A14 eigentune Var1 is evaluated. The comparison is performed using ATLAS data unfolded to particle level at $\sqrt{s}=7$ TeV for the jet multiplicity (a) and the leading jet momentum (b) (Analysis A [4]) as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
4.3 Initial and final state radiation

In Fig. 9–13 the impact of the ISR/FSR parameter variations (Var2-Var3c of the A14 eigentunes) are shown for different kinematic variables and compared to unfolded data. The impact of the variations on the jet multiplicity is shown in Fig. 9. The impact is largest for the variation Var3c, which corresponds to the ISR variations only, and very small for all other variations. Since the lower jet multiplicities are mainly modeled by the ME, the changes in these bins are small. No significant impact is observed for the momentum of the leading jet. For the momentum of the fifth jet and the reconstructed hadronic top quark the impact is of similar size for all variations, while for the reconstructed $t\bar{t}$ momentum the impact is largest for Var3b and Var3c.
Figure 9: The impact of the four Powheg+Pythia8 A14 eigentunes corresponding to initial- and final-state radiation are compared for the jet multiplicity in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis A [4]). The data and generator predictions are represented in the same way as in Fig. 1.
Figure 10: The impact of the four Powheg+Pythia8 A14 eigentunes corresponding to initial- and final-state radiation are compared for the leading jet $p_T$ in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis A [4]). The data and generator predictions are represented in the same way as in Fig. 1.
Figure 11: The impact of the four Powheg+Pythia8 A14 eigentunes corresponding to initial- and final-state radiation are compared for the fifth jet $p_T$ in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis A [4]). The data and generator predictions are represented in the same way as in Fig. 1.
Figure 12: The impact of the four Powheg+Pythia8 A14 eigentunes corresponding to initial- and final-state radiation are compared for the reconstructed hadronic top-quark $p_T$ in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
Figure 13: The impact of the four Powheg+Pythia8 A14 eigentunes corresponding to initial- and final-state radiation are compared for the $p_T$ of the $t\bar{t}$ system in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
4.4 Full radiation systematics

In Sec. 4.3 it has been established that Var3c is the eigentune parameter variation with the largest impact on the kinematic distributions under study. It however does not include the full radiation uncertainty, since the scales in the ME are unchanged. Therefore the combined variation of the hadronization and factorization scale in the ME with the tune variation is tested in the following. Additional samples are produced to define two settings of the parameters that estimate one sigma variations of the radiation systematics as done for Powheg+Pythia6 in Sec. 3. The two samples compared to the nominal Powheg+Pythia8 sample and the data distribution are produced as follows:

- The factorization and hadronization scales are varied by a factor of 0.5, and, simultaneously the $h_{\text{damp}}$ parameter is increased to 2$m_{\text{top}}$. Furthermore the Var3c up variation from the A14 tune is used.

- The factorization and hadronization scales are varied by a factor of 2.0, while the $h_{\text{damp}}$ parameter is unchanged. Furthermore the Var3c down variation from the A14 tune is used.

The variations cover the uncertainties on the data for the distributions in Fig. 14(a,b,d) while for the hadronic top momentum (Fig. 14(c)) the variations are smaller than the experimental uncertainty. These conclusions are very similar to the ones made for Powheg+Pythia6 in Sec. 3.2. The authors of the A14 tuning note do not provide a recommendation for the radiation systematic. Since the impact of Var3c is always covering the variations 2, 3a and 3b, a possible setup for the evaluation of the systematic using Var3c as discussed here, in combination with variations of the hadronization and factorization scales is proposed, with all uncertainties to be added, in quadrature. An alternative set of samples and uncertainties based on Powheg+Pythia8 as nominal sample is proposed for future analyses as shown in Tab. 4.

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>Samples</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Powheg+Pythia8</td>
<td>N/A</td>
</tr>
<tr>
<td>Hard Scatter Generation</td>
<td>Powheg+Pythia8 vs. aMC@NLO+Pythia8</td>
<td>±</td>
</tr>
<tr>
<td>Parton Shower and Hadronization Model</td>
<td>Powheg+Pythia8 vs. Powheg+Herwig++</td>
<td>±</td>
</tr>
<tr>
<td>Scales and Additional Radiation</td>
<td>Powheg+Pythia8 Variations</td>
<td>Δ</td>
</tr>
<tr>
<td>Interference ($Wt$ and $t\bar{t}$)</td>
<td>Powheg+Pythia8 DR vs. DS Schemes</td>
<td>±</td>
</tr>
</tbody>
</table>

Table 4: A proposal for modeling systematic uncertainties for the $t\bar{t}$ and single top processes as well as the samples used in each case using Powheg+Pythia8 as nominal sample. The symbol Δ denotes the difference in the analysis observables using the simulation from the samples column. The notation ±|Δ| indicates that the full difference is symmetrized and applied to the nominal. Without the absolute value bars, Δ indicates that the signed difference with respect to the nominal for all variations is used to estimate the uncertainty.
Figure 14: The impact of the final-state radiation eigentune A14-Var3c, scale variation and damping variation is evaluated. The comparison is performed for the jet multiplicity (a) and the leading jet momentum (b) (Analysis A [4]) as well as for the hadronic top-quark $p_T$ (c) and the $p_T$ of the $t\bar{t}$ system (d) in $t\bar{t}$ lepton+jets events using ATLAS data unfolded to particle level at $\sqrt{s} = 7$ TeV (Analysis B [9]). The data and generator predictions are represented in the same way as in Fig. 1.
5 Additional considerations

5.1 Powheg+Herwig++

Top-quark simulation for Run II used the Herwig++ generator as a replacement of the older Fortran based Herwig simulation. Analyses using the 2015 dataset are already using Herwig++ in combination with both PowHEG-Box and aMC@NLO. One important implication of the change from Run I to Run II is that the recommendation for systematic uncertainties associated with fragmentation have been changed from comparing PowHEG+PYTHIA6 and PowHEG+HERWIG to comparing PowHEG+PYTHIA6 and PowHEG+HERWIG++. Among the first analyses where this was observed was the early ATLAS top-quark pair production cross-section at 13 TeV [29]. The same analysis technique was used in Run II as was used in Run I [30], but there was a significant change in the size of the fragmentation uncertainty from $\sim 1\%$ to $\sim 5\%$. A detailed study was performed to understand how well the PowHEG+HERWIG++ setup models the data and also the impact of various changes in the parameters on the data/simulation agreement. Figure 15 illustrates four significant deviations of the simulation from the unfolded $\sqrt{s} = 7$ TeV data. First, the top-quark $p_T$ spectrum is too soft at low $p_T$ (Fig. 15a). The modeling of the top-quark $p_T$ at low $p_T$ can be improved by setting the Herwig++ option FinalStateReconOption 1 (abbreviated as Recon 1) which resuffles the momenta of the outgoing partons to preserve the invariant mass of the $t\bar{t}$ system. Second, the $t\bar{t}$ system is too far forward compared with the data (Fig. 15b). This seems to be a generic feature of the CT10 PDF set and is remedied by using NNPDF3.0. Third, there are too many events with very high jet multiplicity (Fig. 15c). These jets are not produced from the matrix element and are formally only accurate to leading logarithmic order so we do not expect them to be well-predicted by the simulation. However, these jets are sensitive to parameters related to the matching of the matrix-element calculation and the parton shower and as such can be tuned to improve the data/simulation agreement. This resulting spectrum should not be regarded as prediction of the model, but the agreement is important for various analyses that rely on observables that are sensitive to the total jet multiplicity. Lastly, as seen in Fig. 15d, the hardest jets in $t\bar{t}$ events are too soft. Despite many parameter changes, including the ones listed above, this mis-modeling of the hard jet $p_T$ remains.
Figure 15: Representative distributions of Powheg+Herwig++ compared with unfolded data at $\sqrt{s} = 7$ TeV (Analysis A and B) to illustrate significant deviations of the simulation from the data.
5.2 Single top $Wt$ production and $t\bar{t}$ interference

At leading order, the matrix elements for $t\bar{t}$ production and single top factorize: the superposition of the leading order simulation of $t\bar{t}$ and single top is equivalent to the leading order simulation of the superposition of $t\bar{t}$ and single top. However, at NLO accuracy, there is non-trivial interference. For example, the reaction $gg \rightarrow t\bar{t}^* \rightarrow t\bar{b}W^*$ contributes at LO to $t\bar{t}$ and at NLO to single top. When $m(bW^*) \sim m_t$, this interference is large. The interference is treated by removing a contribution from the $Wt$ simulation using either Diagram Removal (DR) or Diagram Subtraction (DS) schemes (the difference is an estimate of the uncertainty). When the interference between NLO single top and LO $t\bar{t}$ is small, these two schemes are comparable [31, 32, 33]. However, there are many searches for physics beyond the SM at the LHC that select single top events with high purity that have kinematic properties such that the interference with LO $t\bar{t}$ may not be small. The reason for a high $Wt$ purity is that such events are sufficiently different from $t\bar{t}$ events that they are not as suppressed by kinematic requirements aimed at reducing top-quark pair production. However, such searches often require implicitly or explicitly several $b$-tagged jets and other kinematic selections that increase the interference. In such cases, using separated $t\bar{t}$ and $Wt$ processes simulated at NLO may not be meaningful and the difference between the DR and DS interference removal schemes may not give an accurate estimate of the uncertainty on the modeling of the composite process. Specific examples at particle level are shown in Fig. 16 and Fig. 17 for a single lepton and a dilepton selection, respectively. To mimic searches with high missing transverse momentum from unmeasured particles such as supersymmetric neutralinos, events are required to have a magnitude of the momentum from all neutrinos ($E_T^{miss}$) to be at least 200 GeV. Furthermore, to enhance the interference, events are required to have at least four jets with $p_T > 25$ GeV and at least two such jets must have originated from $b$-quarks as described in Sec. 2. For both the single and dilepton selections, there is a significant and increasing difference between the single top simulations with the DR and DS schemes for the $E_T^{miss}$ (Fig. 16) and the leading jet $p_T$ distributions (Fig. 17). In both cases, the DR scheme results in both a harder $E_T^{miss}$ and leading jet $p_T$ spectrum. For this selection the $Wt$ contribution is a small fraction (5–10%) of the total top-quark contribution. However, there are many searches where $Wt$ can be a significant fraction of the explicit $t\bar{t}$ simulation. For example, an ATLAS search for $t\bar{t} + E_T^{miss}$ in the single lepton channel [34] has developed and utilized a series of powerful techniques to suppress the $t\bar{t}$ background, leaving a larger contribution of $Wt$ events. The remaining events are dominated by events with two real leptons, one of which is not reconstructed. This is the reason for the effect in the right plots of Fig. 16 and 17, for which the $Wt$ simulation is nearly 50% of the explicit $t\bar{t}$ simulation in the high $E_T^{miss}$ tail. The 50% difference between the single top DR and DS results in 15–20% differences in the combined $t\bar{t} + Wt$ simulation when comparing the two interference schemes. Higher $Wt$ purities are possible and so it is critical to understand the validity of this approach to estimate the systematic uncertainty.

In Run I, the interference in this extreme region of phase space was further studied by looking at a LO sample generated with AcerMC [35, 36, 37] with the inclusive $2 \rightarrow 6$ reaction $pp \rightarrow W^+W^-b\bar{b}$ that includes the double resonant $t\bar{t}$ production, the single resonant $Wt$ production in association with a $b$-quark and the non-resonant diboson production in association with jets [34]. The analogous study in Run II uses MadGraph5_aMC@NLO interfaced with Pythia8. Figures 18 and 19 include this inclusive $WWbb$ sample in comparison with the Powheg sample using the DR scheme. There are significant differences in both the $E_T^{miss}$ and leading jet $p_T$ distributions. This is due in part to the fact that the $t\bar{t}$ component of the $WWbb$ sample is LO and there are significant NLO corrections. Further studies of the full ME comparison will benefit from a multileg $WWbb$ simulation, which requires a non-trivial merging setup, and ultimately a full NLO $WWbb$ simulation.
Figure 16: The distribution of the magnitude of the momentum from all neutrinos ($E_{\text{miss}}$) for events passing a one lepton (a) and a two lepton (b) selection. Both selections require at least four jets with $p_T > 25$ GeV, of which two must have originated from $b$-quarks. All distributions are normalized to have the same integral in the above range. The gray band in the ratio is the statistical uncertainty from the simulation using the DR scheme and the uncertainty on the markers is from the simulation used in the numerator of the ratio. Most of these uncertainties are smaller than the markers.

Figure 17: The distribution of the leading jet $p_T$ passing a one lepton (a) and a two lepton (b) selection. Both selections require $E_{\text{miss}} > 200$ GeV and at least four jets with $p_T > 25$ GeV, of which two must have originated from $b$-quarks. All distributions are normalized to have the same integral in the above range. The gray band in the ratio is the statistical uncertainty from the simulation using the DR scheme and the uncertainty on the markers is from the simulation used in the numerator of the ratio. Most of these uncertainties are smaller than the markers.
Figure 18: The distribution of the magnitude of the momentum from all neutrinos ($E_{\text{miss}}$) for events passing a one lepton (a) and a two lepton (b) selection. Both selections require at least four jets with $p_T > 25$ GeV, of which two must have originated from $b$-quarks. All distributions are normalized to have the same integral in the above range. The gray band in the ratio is the statistical uncertainty from the simulation using the DR scheme and the uncertainty on the markers is from the simulation used in the numerator of the ratio. Most of these uncertainties are smaller than the markers.

Figure 19: The distribution of the leading jet $p_T$ passing a one lepton (a) and a two lepton (b) selection. Both selections require $E_{\text{miss}} > 200$ GeV and at least four jets with $p_T > 25$ GeV, of which two must have originated from $b$-quarks. All distributions are normalized to have the same integral in the above range. The gray band in the ratio is the statistical uncertainty from the simulation using the DR scheme and the uncertainty on the markers is from the simulation used in the numerator of the ratio. Most of these uncertainties are smaller than the markers.
6 Summary

Various MC generators have been compared to ATLAS measurements of top-quark pair production. Recommendations for the evaluation of $t\bar{t}$ and single top modeling uncertainties have been presented and the different generator setups were directly compared to each other. The PowHEG+HERWIG++ sample has more jets and a much softer $p_T$ spectrum than the other generators. This leads to a worse description of the data than the PowHEG+PYTHIA6 sample. Comparison of alternative generator setups based on aMC@NLO show smaller effects.

The usage of EvtGen has been shown to lead to generator independent $b$-tagging efficiencies. In addition, a new PowHEG+PYTHIA8 setup has been presented that leads to an improved description of data. Based on this setup, the radiation systematics taking into account scale variations, damping variations and the A14 eigentune variations have been studied. A first possible setup for this systematic was proposed.
References


