



ATLAS PUB Note

ATL-PHYS-PUB-2025-036

17th September 2025



Towards a Systematic Uncertainty Model for the POWHEG $bb4\ell$ Monte Carlo sample

The ATLAS Collaboration

This note presents a study of the ATLAS Monte Carlo (MC) sample describing $pp \rightarrow b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$ events at next-to-leading order in quantum chromodynamics. Events are generated using the POWHEG Box RES implementation of the $bb4\ell$ process and are interfaced with PYTHIA 8 for parton showering and hadronization. The nominal ATLAS $bb4\ell$ MC sample is compared to predictions from POWHEG+PYTHIA 8 for top-quark pair production ($t\bar{t}$) and single-top production in association with a W -boson (tW). The $bb4\ell$ MC sample is intended to reduce modelling uncertainties in precision top-quark physics analyses with respect to the modelling uncertainties associated with the $t\bar{t}+tW$ MC predictions. The impact of theoretical modeling variations is evaluated and compared to the corresponding uncertainty estimate in the $t\bar{t}$ simulation. The studied uncertainties include matrix-element scale variations, changes in the POWHEG+PYTHIA 8 matching parameters, modification of the PYTHIA 8 final-state splitting kernel via μ_R variations, the use of parton distribution function (PDF) sets with an alternative value of $\alpha_S(m_Z)$ and the inclusion of inverse width corrections in $bb4\ell$. These studies support the development of a systematic uncertainty model for the $bb4\ell$ simulation.

1 Overview

In this document, a collection of plots is presented, which is used to study the systematic modelling uncertainty prescription associated with the nominal ATLAS $bb4\ell$ Monte Carlo (MC) sample. The $bb4\ell$ generator [1, 2] describes the full $pp \rightarrow b\bar{b}\ell^+\nu_\ell\ell'^-\bar{\nu}_{\ell'}$ process at next-to-leading order (NLO) in quantum chromodynamics (QCD), thus incorporating top-quark off-shell effects, interference between $t\bar{t}$ and tW production, top-quark decays, and spin correlations at NLO accuracy. In the $bb4\ell$ process version used for the ATLAS MC sample production, a code modification¹ in the finalisation of the hard scatter event is applied to allow also same lepton-flavour combinations $\ell = \ell'$.

Several ATLAS top-quark analyses are limited by the modelling uncertainties assigned due to the approximate treatment of the top-quark off-shell effects and the interference between $t\bar{t}$ and tW production in the nominal $t\bar{t} + tW$ ATLAS predictions, for example the recent top-quark entanglement analysis [3], the $t\bar{t}$ production threshold analysis [4] or analyses using kinematic end-points, as the $m_{b\ell}^{\text{minimax}}$ distribution, which was unfolded in Ref. [5].

The nominal $bb4\ell$ ATLAS MC sample is compared to the nominal $t\bar{t} + tW$ ATLAS MC samples, describing separately the $t\bar{t}$ and tW production process. In the nominal ATLAS $t\bar{t}$ MC sample, generated using the POWHEG method [6, 7], the decays of the on-shell $t\bar{t}$ pair are performed within POWHEG, which also includes a Breit-Wigner smearing of the top-quark invariant mass to approximate off-shell effects. The nominal ATLAS tW prediction approximates the interference between $t\bar{t}$ and tW production with the diagram removal (DR) scheme [8]. In addition to the nominal $t\bar{t} + tW$ prediction, two alternative MC samples are typically considered to estimate uncertainties arising from the treatment of top-quark off-shell effects and $t\bar{t}/tW$ interference. The first alternative MC sample models top-quark decays using MADSPIN rather than the internal POWHEG implementation [9], while the second compares the nominal DR scheme to the diagram subtraction scheme (DS) approach for $t\bar{t} + tW$ interference [10]. In contrast, these comparisons are not needed if the $bb4\ell$ prediction is used.

The $bb4\ell$, $t\bar{t}$ and tW processes are generated using the POWHEG method at a center of mass energy of $\sqrt{s} = 13$ TeV. The $bb4\ell$ process [1, 2] is implemented in the POWHEG BOX RES framework [11]², while the $t\bar{t}$ and the tW processes are simulated using the POWHEG BOX v2 framework [12]. Specifically, the $t\bar{t}$ process uses the hvq implementation [13], and the tW process is modelled with ST_wtch_DR [14].

Depending on the final-state kinematics, each POWHEG generated $bb4\ell$ hard-scatter event has a resonance history assigned. In the $bb4\ell$ -dl (beta) version used in this study, three possible resonance histories are available: a $t\bar{t}$, $tW\bar{b}$ and $\bar{t}Wb$ resonance history [2]. For the $bb4\ell$ generator, different central scale choices (μ_0) for the renormalisation (μ_R) and factorisation scale (μ_F) are applied depending on the resonance history:

- For $bb4\ell$ events, which are assigned a $t\bar{t}$ resonance history,

$$\mu_0 = [(m_t^2 + p_{T,t}^2)(m_{\bar{t}}^2 + p_{T,\bar{t}}^2)]^{1/4} \quad (1)$$

¹ The corresponding code changes can be found in https://gitlab.cern.ch/tjezo/powheg-box-res-bb4l-sl-beta/-/merge_requests/12.

² For the $bb4\ell$ MC sample generation, the allrad emission generation scheme is employed, which allows to generate multiple POWHEG emissions per event: from the production and from the resonance decay processes. This setting is not available for the on-shell generators ($t\bar{t}$ and tW MC samples) discussed in this note, since only the production process is part of the matrix-element calculation.

- For $bb4\ell$ events, which are assigned a tW resonance history ($tW\bar{b}$ or $\bar{t}Wb$),

$$\mu_0 = [(m_{t/\bar{t}}^2 + p_{T,t/\bar{t}}^2)(m_{\bar{b}/b}^2 + p_{T,\bar{b}/b}^2)]^{1/4}, \quad (2)$$

with the transverse momentum p_T and the invariant mass m of the top- and anti-top quark, as well as bottom and anti-bottom quark at the underlying Born level, meaning before the real emission in the POWHEG framework.

In the $t\bar{t}$ hvq simulation, the central scale is defined as

$$\mu_0 = \sqrt{p_{T,t}^2 + m_{\text{top}}^2}, \quad (3)$$

where m_{top} is the invariant mass of the top-quark in the ME calculation, which is always equal to the nominal value of $m_{\text{top}} = 172.5$ GeV, since on-shell top-quarks are described in ME calculation. The transverse momentum of the top-quark $p_{T,t}$ is also evaluated at the underlying Born level. For the tW process [14], also a dynamic scale choice is applied [10], given by

$$\mu_0 = H_T/2, \text{ with } H_T = \sqrt{m_t^2 + p_{T,t}^2} + \sqrt{m_W^2 + p_{T,W}^2} (+p_T^j). \quad (4)$$

The transverse momentum of an additional jet p_T^j is included in the H_T definition, if a real emission was generated in the POWHEG framework.

In the POWHEG generation of these MC samples, the NNPDF3.0NLO [15] ($\alpha_S(m_Z) = 0.118$) PDF set is used and the top-quark mass is set to $m_{\text{top}} = 172.5$ GeV. The POWHEG h_{damp} parameter is set to $h_{\text{damp}} = 1.5m_{\text{top}}$. While the $t\bar{t}$ and tW predictions are obtained with a five flavour number scheme (5FNS) calculation, the $bb4\ell$ calculation is done within the 4FNS. Thereby, matching factors are applied to match the 4FNS matrix element calculation to the 5FNS PDF set [2].

All POWHEG generated Les Houches events (LHEs) are interfaced with PYTHIA 8 [16, 17] in the nominal MC samples for modelling of the parton shower and hadronisation, thereby using the A14 tune [18]. The PYTHIA 8 versions of the considered samples differ between the $bb4\ell$ (8.312)³, the nominal hvq $t\bar{t}$ ATLAS sample (8.230), the hvq +MADSPIN +PYTHIA 8 $t\bar{t}$ sample (8.309) and the tW DR (8.309) and DS (8.307) samples. It was verified that the usage of different PYTHIA 8 versions has a minor impact on the presented distributions. The POWHEG-PYTHIA 8 matching parameter p_T^{def} is set to 2 for all described MC samples. While the default minimum POWHEG emission generation scale ptsqmin is set to 0.8 GeV^2 in case of the $t\bar{t}$ and tW MC samples, in case of the $bb4\ell$ sample it is instead set to 1.44 GeV^2 ⁴.

Different PYTHIA 8 recoil strategies are used in the nominal ATLAS $t\bar{t} + tW$ and $bb4\ell$ MC samples, which affect the emission pattern of the second and subsequent parton shower emissions from top-quark decay products. While for the nominal ATLAS $t\bar{t} + tW$ MC sample the recoil option recoil-to-colour is used, in the case of the nominal $bb4\ell$ sample instead the recoil-to-top option is applied, which should give rise

³ A source code change was included in the PYTHIA 8.312 version in ATLAS, which enables the recoil-to-top setting for $bb4\ell$ events and which is part of the official PYTHIA 8.313 release.

⁴ This parameter was reset in the $bb4\ell$ MC sample production in order to have a more comparable set-up between the PYTHIA 8 and HERWIG 7 matched hard-scatter events, since the tuned parton shower cut-off scale for the previous HERWIG 7.2 version of 0.958 GeV is higher than the minimum p_T generated by POWHEG. This is not the case though for the newer HERWIG 7.3 version with a shower-cut off scale of 0.655 GeV [19]. A comparison of the nominal with an alternative $bb4\ell$ MC sample, in which the default value of $\text{ptsqmin} = 0.8 \text{ GeV}^2$ was used, showed that the parameter setting does not have a large influence on the considered differential distributions.

to wider-angle gluon radiation [20]. To allow a more direct comparison between setups an alternative $t\bar{t}$ POWHEG+PYTHIA 8 MC sample with the recoil-to-top setting and PYTHIA version 8.245 is also shown.

For all processes the EVTGEN program [21] was used to simulate the decay of the bottom and charms hadrons. All distributions are normalised to their NLO+PS cross-section, which is calculated during the event generation.

Two aspects of the $bb4\ell$ MC generation are studied with dedicated MC samples:

- Inverse width correction [2]: In narrow-width approximation (NWA) calculations, top-quark production and decay are treated as separate processes, as is the case in the $t\bar{t}$ and tW MC sample generation. When integrating over the decay phase space of the full differential cross section, $d\sigma_{\text{prod}\times\text{dec}}$, one should recover the production cross section $d\sigma_{\text{prod}}$ to all orders of perturbation theory. However, an additional contribution — known as the spurious width term — arises after integration. This term is typically avoided in NWA calculations by using a perturbative expansion of the inverse decay width $1/\Gamma_{\text{NLO}}$. A similar correction can also be applied in the context of full off-shell calculations [2]. The use of these inverse-width or spurious-width corrections is particularly relevant in the comparison of the off-shell $bb4\ell$ calculation with the $t\bar{t} + tW$ predictions, since genuine effects from top-quark off-shell modeling and $t\bar{t}/tW$ interference can be disentangled from spurious width effects. The nominal ATLAS $bb4\ell$ MC sample is thereby defined without applying the inverse-width corrections. The variation when applying the inverse width correction is evaluated via event-weights in the nominal $bb4\ell$ sample. Further, in Ref. [2] it is argued that the spurious width corrections should follow a similar kinematic dependence as the NLO/LO K -factor in case of narrow-width-approximation calculations. To study this in case of the $bb4\ell$ calculation, an additional $bb4\ell$ LO QCD sample was produced, which was matched to PYTHIA 8.312 using the A14 tune without applying any NLO parton shower matching procedure and using the default PYTHIA 8 settings `Space/TimeShower:pTmaxMatch=0/1`.
- Dedicated $bb4\ell$ matching: A dedicated $bb4\ell$ parton shower matching procedure⁵ is required, since up to three emissions per hard-scatter event can be generated by PowHEG in the allrad scheme used in $bb4\ell$: an emission from the production, the top-quark and the anti-top quark decay. Consequently, up to three matching scales have to be respected during the parton showering. The nominal $bb4\ell$ +PYTHIA 8 and a $bb4\ell$ +HERWIG 7.3.0 MC sample are compared to $bb4\ell$ +PYTHIA 8 and $bb4\ell$ +HERWIG 7 MC samples, which apply the standard PowHEG-PYTHIA or PowHEG-HERWIG matching as done e.g. in the $t\bar{t}$ hvq MC sample. In the standard PowHEG matching procedure only the matching scale associated with the PowHEG emission from the production process is respected.

To study modelling uncertainties in the ATLAS $bb4\ell$ MC sample and to compare to the corresponding uncertainties in the $t\bar{t} + tW$ or the $t\bar{t}$ MC sample, the following predictions are used:

- Matrix-element (ME) μ_R and μ_F variations: The ME scale variation uncertainty is evaluated via event-weights in the nominal $bb4\ell$, $t\bar{t}$ and tW DR MC sample.
- PYTHIA 8 splitting kernel variations [23]: The PYTHIA 8 splitting kernel variation uncertainty is evaluated via event-weights in the nominal $bb4\ell$ MC sample. A $t\bar{t}$ hvq +PYTHIA 8.307 MC sample was produced with the same PowHEG and PYTHIA settings as for the $t\bar{t}$ nominal MC sample, but also including splitting kernel variation event weights.

⁵ The dedicated $bb4\ell$ matching is implemented through a $bb4\ell$ UserHook in PYTHIA 8 and a dedicated shared library in HERWIG 7, similar to the implementation described in Ref. [22].

- Matching uncertainty: Additional MC samples are generated for both $bb4\ell$ and $t\bar{t}$ hvq by only changing the p_T^{hard} POWHEG+PYTHIA 8 matching parameter from the default value of 0, to the alternative values of 1 and 2 [9]. Thereby, in case of the $t\bar{t}$ hvq +PYTHIA MC sample, the newer PYTHIA 8.306 version with respect to the nominal $t\bar{t}$ hvq +PYTHIA MC sample is used. Another $bb4\ell$ MC sample is produced to evaluate the parton shower matching uncertainty in the top-quark decay [22]. In this variation, effectively the definition of the transverse momentum (p_T) of parton shower emissions originating from the top-quark decay is changed from the POWHEG p_T definition used in the top-quark decay in $bb4\ell$ (default) to the PYTHIA p_T definition⁶. The p_T definition of a parton shower emission influences the parton shower matching, since this p_T -scale is compared to the matching scale.
- α_S in POWHEG Sudakov: For both processes, $bb4\ell$ and hvq, two new MC samples generated with the NNPDF3.0 PDF set with $\alpha_S(m_Z) = 0.115$ (LHAPDF ID 264000) and $\alpha_S(m_Z) = 0.121$ (LHAPDF ID 267000) are produced as suggested in [22], leaving the other MC generator settings unchanged and using the PYTHIA 8.312 version.

The histograms used to study the $bb4\ell$ systematic uncertainty model are produced using RIVET [24].

To study properties of the particle-level b -jets, the publicly available RIVET routine MC_HFDECAYS [25] is used, thereby showing the

- b -jet width, defined as the sum of the p_T weighted distance of the particles in a jet to the jet-axis, normalised to the p_T sum of all particles in the jet

$$\frac{\sum_{i \in \text{jet}} p_T^i \cdot \Delta R(i, \text{jet-axis})}{\sum_{i \in \text{jet}} p_T^i}. \quad (5)$$

- the b -jet fragmentation function, defined as the fraction of the total three-momentum of the jet carried by the b -hadron in direction of the jet $\vec{p}_{\text{had}} \cdot \vec{p}_{\text{jet}} / \vec{p}_{\text{jet}}^2$.

Additionally, a RIVET routine targeting a typical $t\bar{t}$ phase space is considered. Dressed particle-level leptons - obtained by adding prompt photons to the bare prompt lepton in a cone of $\Delta R = 0.1$ - are required to satisfy the transverse momentum cut $p_T > 20 \text{ GeV}$ and $|\eta| < 2.5$. Jets are defined by means of the anti- k_t clustering algorithm [26] with the radius parameter $R = 0.4$ using the FastJet implementation [27]. The particle level jets are required to satisfy the kinematic cuts $p_T > 25 \text{ GeV}$ and $|\eta| < 2.5$. In the jet clustering algorithm all stable particles with a mean lifetime greater than $0.3 \times 10^{-10} \text{ s}$ are considered, except for the dressed prompt leptons and prompt neutrinos. An overlap removal procedure is applied at particle level: lepton objects which are close to a reconstructed jet ($\Delta R(\text{jet}, \text{lepton}) < 0.4$) are excluded to obtain particle level selections which are more similar to typical detector level object selections. A jet is identified as a b -jet, if a b -hadron with $p_T > 5 \text{ GeV}$ is associated to the jet through the ghost-matching technique from Ref. [28]. In the event selection, exactly two charged leptons are required (either ee , $\mu\mu$ or $e\mu$), which need to have opposite electric charge. Additionally, at least two b -jets need to be present in the event. The top-quarks at particle level are reconstructed using the two hardest truth neutrinos from the event record. The pairing of the two neutrinos and the two charged, dressed leptons that originate from the same W decay is found by minimizing $|m_{\ell_i \nu_i} - 80.4 \text{ GeV}| + |m_{\ell_j \nu_j} - 80.4 \text{ GeV}|$. Then, the particle level top quarks are constructed, by finding Wb -pairs which minimize the difference between m_{Wb} and $m_{\text{top}} = 172.5 \text{ GeV}$. The following observables are considered:

⁶ This is achieved by setting `ScaleResonance:veto=1` and `FSREmission:veto=0` in the $bb4\ell$ PYTHIA 8 UserHook.

- Invariant mass of the two ℓb -pairs per event, which are closest in distance in ΔR - thereby first combining the hardest b -jet with the closest charged lepton and requiring exactly two b -jets ($N_{b\text{-jet}} = 2$)
- Invariant mass of the reconstructed top-quark, which is defined as the invariant mass of the $\ell \nu b$ -pair with the positively charged lepton ℓ^+
- Transverse momentum distribution of the second-hardest light jet, i.e. a particle level jet, to which neither a b -hadron nor a c -hadron has been ghost-matched.

Additionally, without applying any event selection, one parton-level distribution is considered:

- Invariant mass of the last parton-level top-quark before the top-quark decay in the event record. This definition does not include the parton-level anti-top quark.

In Figure 1(a), the invariant mass distribution of the reconstructed top-quark m_t^{reco} is shown, in which the m_t^{reco} -peak is sensitive to the PYTHIA 8 recoil-setting, while the influence of the top-quark lineshape uncertainty (evaluated with the alternative hvq + MADSPIN + PYTHIA 8 MC sample) is visible in the full reconstructed invariant mass range. In Figure 1(b), the invariant mass distribution of the ℓb -pair is shown, in which the uncertainty in the $t\bar{t} + tW$ prediction for the high $m_{\ell b}$ -tail is completely dominated by the $t\bar{t}/tW$ -interference uncertainty - evaluated by comparing the $t\bar{t} + tW$ DR to $t\bar{t} + tW$ DS prediction.

In Figure 2, the same $m_{\ell b}$ -distribution is displayed, but showing the distributions obtained with the $t\bar{t}$, $t\bar{t} + tW$, $bb4\ell$ and only $bb4\ell$ events which were projected onto a $t\bar{t}$ resonance history. It is observed, that in the high $m_{\ell b}$ phase-space region, the contribution of events from the tW sample and from $bb4\ell$ events projected onto a tW resonance history has increased importance.

Thus, also when regarding the matrix-element scale variations in the $m_{\ell b}$ -distribution - combined or separately in the $t\bar{t}$ and tW like events - in Figure 3, the importance of the ME scale variations in the tW events and $bb4\ell$ events projected onto a tW topology increases for high $m_{\ell b}$ - albeit still remaining small.

In the b -jet property distributions shown in Figure 4 - the b -jet fragmentation function and the b -jet width - the dedicated $bb4\ell$ matching has a similar effect for both the PYTHIA and the HERWIG parton shower, thereby validating the two independent implementations.

In Figure 5, the transverse momentum distribution of the second hardest light jet is shown, and its variation when applying the p_T^{hard} matching uncertainty definition, which is in very good agreement between the $bb4\ell$ and hvq sample. The b -jet fragmentation functions obtained with the nominal $bb4\ell$ MC sample and its variation when applying the matching uncertainty in top-quark decay is shown in Figure 6. The matching uncertainty in top-quark decays is only applicable in $bb4\ell$ and was found to influence b -jet properties.

In Figure 7, the b -jet fragmentation function is shown, as predicted by the $bb4\ell$ and $t\bar{t} + tW$ MC samples, as well as showing the variation induced by the PYTHIA 8 splitting kernel variations and by using a different $\alpha_S(m_Z)$ value in the POWHEG Sudakov. Additionally, the b -jet fragmentation function prediction of the $t\bar{t} + tW$ MC samples is shown, using the $t\bar{t}$ hvq MC sample with the recoil-to-top PYTHIA 8 setting. This comparison emphasizes that the $bb4\ell$ and $t\bar{t} + tW$ b -jet fragmentation function predictions are in much better agreement when using the same recoil-setting, as would be expected due to the large influence of this parton shower setting on the b -jet fragmentation function. The μ_R variation in the PYTHIA 8 FSR heavy quark gluon splitting kernels $Q \rightarrow Qg$ ($Q \in \{t, b\}$) are the dominant splitting kernel variations in the b -jet fragmentation function and the corresponding variation is significantly larger in case of the $t\bar{t}$

$h\nu q + \text{PYTHIA 8}$ compared to the $bb4\ell + \text{PYTHIA 8}$ prediction. Instead, when varying the $\alpha_S(m_Z)$ value in the PowHEG Sudakov by using an alternative PDF set with a different $\alpha_S(m_Z)$ value, only a variation in the $bb4\ell$, but not in the $t\bar{t} h\nu q$ prediction is observed. Thereby, the varied samples generated with an alternative PDF set are compared to the prediction, which is obtained by reweighting the nominal predictions to the same alternative PDF set. This event weight reweighting does not affect the PowHEG Sudakov though, such that the above comparison only focuses on the variation of $\alpha_S(m_Z)$ in the PowHEG Sudakov, and not in the matrix element calculation of the so-called underlying Born configuration, which is varied in both cases: when using the predictions obtained by explicitly setting an alternative PDF set or using event weight reweighting. The difference in the PowHEG $\alpha_S(m_Z)$ variation uncertainty can be explained by the fact that in the $t\bar{t} h\nu q$ MC sample, the first gluon from a top-quark decay is always generated by the PYTHIA 8 parton shower, making it sensitive to variations of the FSR splitting kernels. In the $bb4\ell$ MC sample, the first gluon emission from a top-quark decay can be generated in the PowHEG event generation, rendering it less sensitive to PYTHIA 8 variations but sensitive to the $\alpha_S(m_Z)$ value in the PowHEG Sudakov factor. The variation in the first gluon emission from the top-quark decay would thus partly be shifted from the PYTHIA 8 splitting kernel variations to the variations in the PowHEG Sudakov factor in case of the $bb4\ell$ simulation. This interpretation seems also plausible since the shape of the FSR μ_R variation in the PYTHIA 8 $Q \rightarrow Qg$ splitting kernel and the $\alpha_S(m_Z)$ in the PowHEG Sudakov factor in the b -jet fragmentation function is very similar, which would be expected if both variations describe the same physical effect in the first emission from the top-quark decay, which can take place either in the PYTHIA 8 or in the PowHEG framework in case of the $bb4\ell$ simulation.

The invariant mass distribution of the last parton-level top-quark in the event record and of the reconstructed top-quark are shown in Figure 8. While the inverse width correction introduces a significant normalisation and shape effect in case of the invariant mass of the parton-level top-quark - where the shape variation follows the NLO/LO $bb4\ell$ K-factor - for the reconstructed top-quark mass at particle level the shape variation is strongly reduced. The variation in the overall normalisation due to the inverse width correction was observed already in Ref. [2].

In conclusion, the $bb4\ell$ MC sample removes the need for the top-quark off-shell effects and $t\bar{t}/tW$ interference uncertainties that are applied to the $t\bar{t} + tW$ MC prediction, benefiting top-quark analyses, which are sensitive to these modelling uncertainties. Systematic uncertainties evaluated with the same prescription in $bb4\ell$ and $t\bar{t} + tW$ - as e.g. the matching uncertainty in the production process - lead to similar uncertainty estimates in both predictions. While the usage of the multiple emission generation scheme in $bb4\ell$ leads to dedicated uncertainties such as the matching uncertainty in top-quark decay and the $\alpha_S(m_Z)$ variation in the PowHEG Sudakov factor - which are not present in the $t\bar{t} + tW$ modelling uncertainty recipe - the overall modelling uncertainties are expected to be reduced in the $bb4\ell$ compared to $t\bar{t} + tW$ predictions.

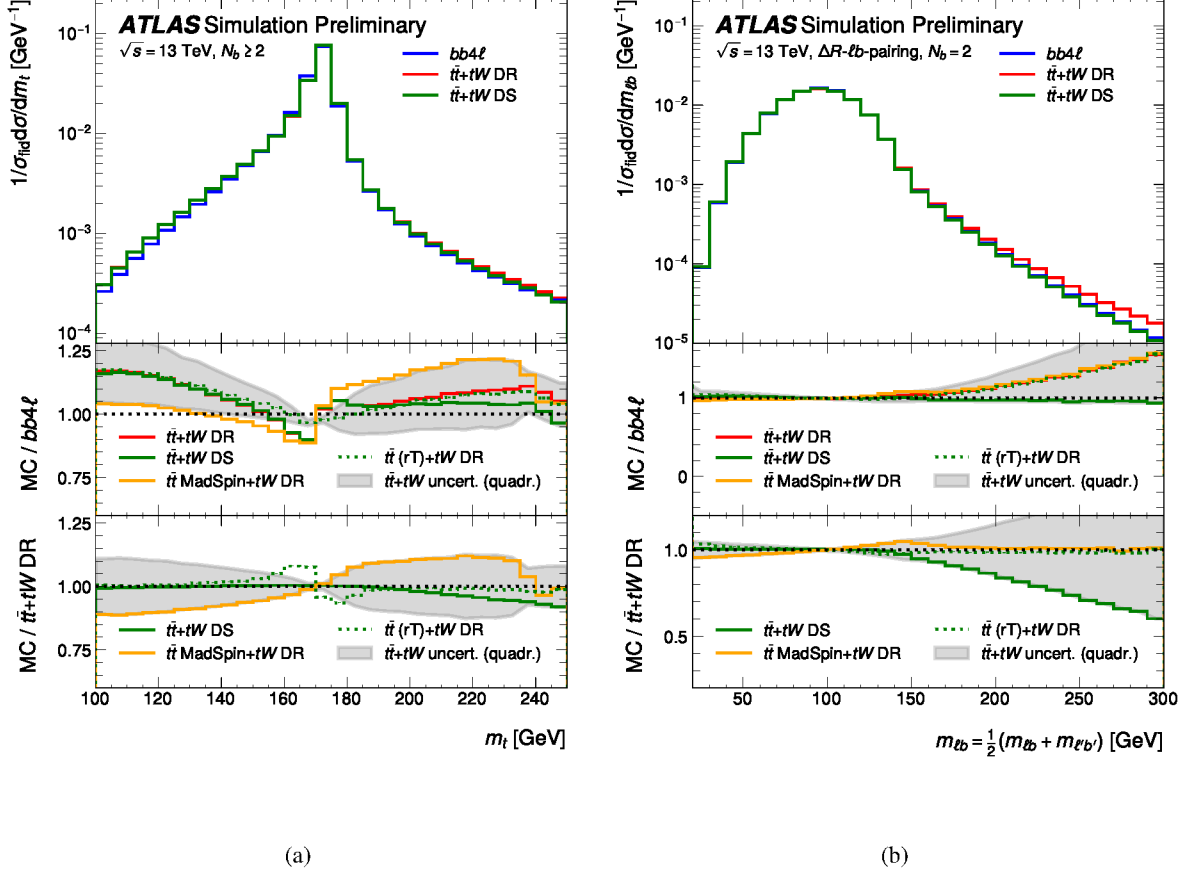


Figure 1: Comparison of the predictions for (a) the invariant mass distribution of the reconstructed top-quark (identified through the positively charged lepton in the $b\ell\nu$ pair) and for (b) the invariant mass distribution of the ℓb -pair obtained through ΔR -pairing of the nominal ATLAS $bb4\ell$ and $\tilde{t}\tilde{t}+tW$ DR MC sample. Also, the comparison to $\tilde{t}\tilde{t}+tW$ MC samples is included, which are used to evaluate the systematic uncertainties related to the $\tilde{t}\tilde{t}/tW$ interference ($\tilde{t}\tilde{t}+tW$ DS), the top-quark line shape modelling ($\tilde{t}\tilde{t} \text{ hvq} + \text{MADSPIN} + \text{PYTHIA 8} + tW \text{ DR}$) and the PYTHIA 8 recoil-scheme ($\tilde{t}\tilde{t} \text{ hvq} + \text{PYTHIA 8 recoil-to-top} + tW \text{ DR} = \tilde{t}\tilde{t} (rT) + tW \text{ DR}$). The gray uncertainty band is obtained by adding in quadrature the uncertainty estimates obtained with the lineshape and the interference uncertainty w.r.t. to the nominal $\tilde{t}\tilde{t}+tW$ DR ATLAS prediction and symmetrising the uncertainty. In the middle panel, this uncertainty band is applied to the $\tilde{t}\tilde{t} (rT) + tW \text{ DR}$ prediction to be able to compare better to the $bb4\ell$ prediction, which is produced with the same PYTHIA 8 recoil scheme. In the bottom panel instead this uncertainty band is applied to the nominal ATLAS $\tilde{t}\tilde{t}+tW \text{ DR}$ prediction. The differential distributions are normalised to the fiducial phase space, i.e. to the cross-section of all events passing the particle-level event selection. When using the $bb4\ell$ prediction as a nominal MC sample the $\tilde{t}\tilde{t}/tW$ interference and the top-quark line shape uncertainty do not need to be applied, since these effects are described at NLO QCD accuracy in the $bb4\ell$ ME calculation. This leads to a significant improvement in the modelling uncertainties in the invariant mass distribution of the reconstructed top-quark, in which the line shape uncertainty is significant and the high $m_{\ell b}$ -tail, in which the DR/DS interference uncertainty is dominant.

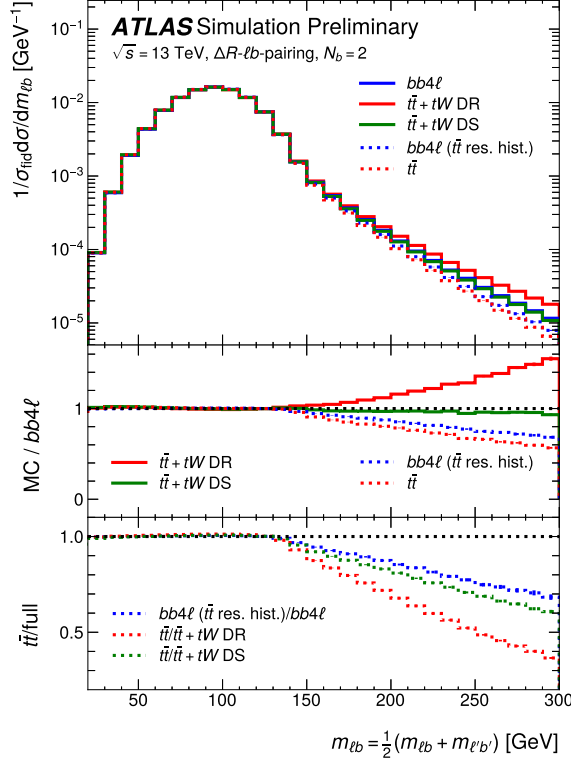


Figure 2: Comparison of the predictions for the invariant mass distribution of the ℓb -pair obtained through ΔR -pairing of the nominal ATLAS $bb4\ell$, $t\bar{t} + tW$ DR MC sample and the alternative $t\bar{t} + tW$ DS MC sample. Further, the comparison of the prediction obtained only with $t\bar{t}$ events (red dotted) or only with $bb4\ell$ events which were projected onto a $t\bar{t}$ resonance history (blue dotted) is shown. The differential distributions are normalised to the fiducial phase space, i.e. to the cross-section of all events passing the particle-level event selection. The relative importance of tW events in the $t\bar{t} + tW$ prediction and $bb4\ell$ events, which are projected onto a tW resonance history is enhanced in the high $m_{\ell b}$ -tail. This is in good agreement with the observation of an increased DR/DS interference uncertainty (Figure 1(b)) and ME scale variation uncertainty in tW -like events (Figure 3) in this phase-space region. Comparing the fractions of $t\bar{t}$ events — taken either from the $t\bar{t}$ sample in the $t\bar{t} + tW$ prediction or from $bb4\ell$ events projected onto a $t\bar{t}$ resonance history — in $t\bar{t} + tW$ and $bb4\ell$ shows a similar overall trend. However, the $t\bar{t}$ fraction in the $t\bar{t} + tW$ DR prediction decreases more rapidly, while the $t\bar{t} + tW$ DS prediction is in closer agreement with the $t\bar{t}$ fraction in $bb4\ell$.

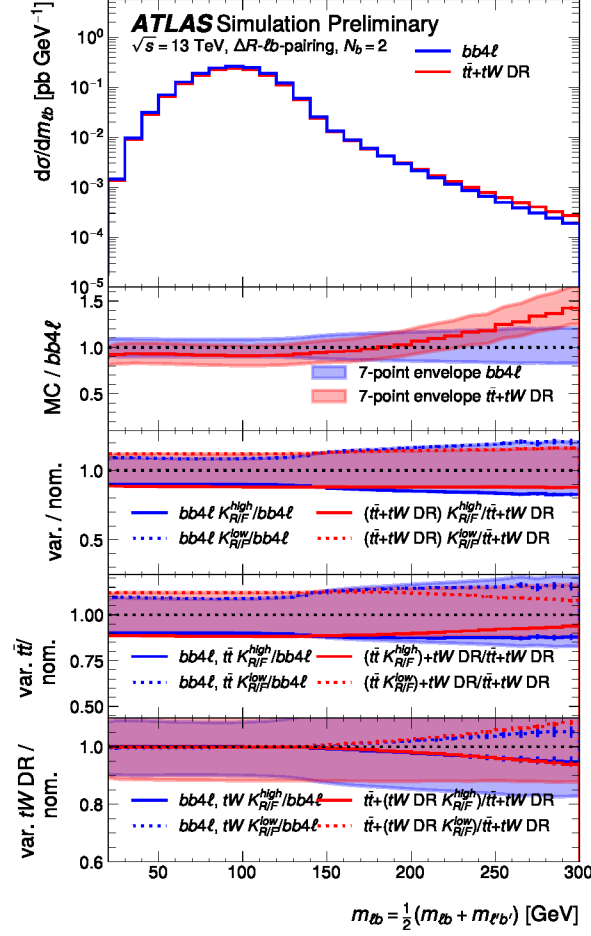


Figure 3: Comparison of the predictions for the invariant mass distribution of the ℓb -pair obtained through ΔR -pairing of the nominal ATLAS $bb4\ell$ and $t\bar{t} + tW$ DR MC sample. The envelope of the seven-point ME scale variation - obtained by varying $\mu_R = K_R\mu_0$ and $\mu_F = K_F\mu_0$ in all $bb4\ell$ events and all $t\bar{t}$ and tW events simultaneously by $(K_R, K_F) \in \{(0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (1, 2), (2, 1), (2, 2)\}$ - is displayed as blue and red band for the $bb4\ell$ and $t\bar{t} + tW$ predictions, respectively, in all four ratio panels. In the second ratio panel, the envelope of the seven-point ME scale variation is compared to varying $K_R = K_F = 0.5$ ($K_{R/F}^{\text{low}}$, dotted) and $K_R = K_F = 2$ ($K_{R/F}^{\text{high}}$, solid) for all $bb4\ell$ events (blue) and all $t\bar{t} + tW$ events (red). It is thereby shown, that the full seven-point ME scale variation is well covered by the three-point ME scale variation. In the third ratio, this three-point ME scale variation uncertainty, i.e. $K_R = K_F = 0.5$ ($K_{R/F}^{\text{low}}$, dotted) and $K_R = K_F = 2$ ($K_{R/F}^{\text{high}}$, solid) is evaluated only for $t\bar{t}$ events in the $t\bar{t} + tW$ simulation stack and $bb4\ell$ events, which are projected onto a $t\bar{t}$ resonance history. In the bottom ratio pad, the same comparison is done, but the three-point ME scale variation applied only onto tW events of the $t\bar{t} + tW$ simulation stack or $bb4\ell$ events, which are projected onto a tW resonance history. The increased importance of the ME scale variation in tW events in the high $m_{\ell b}$ -tail is in good agreement with the increased tW fraction in this phase space region in Figure 2. Further, the ME scale variation in tW events in the $t\bar{t} + tW$ and $bb4\ell$ predictions are in good agreement, even though different central scale definitions are used and the $bb4\ell$ resonance history projectors rely on LO matrix-elements [2].

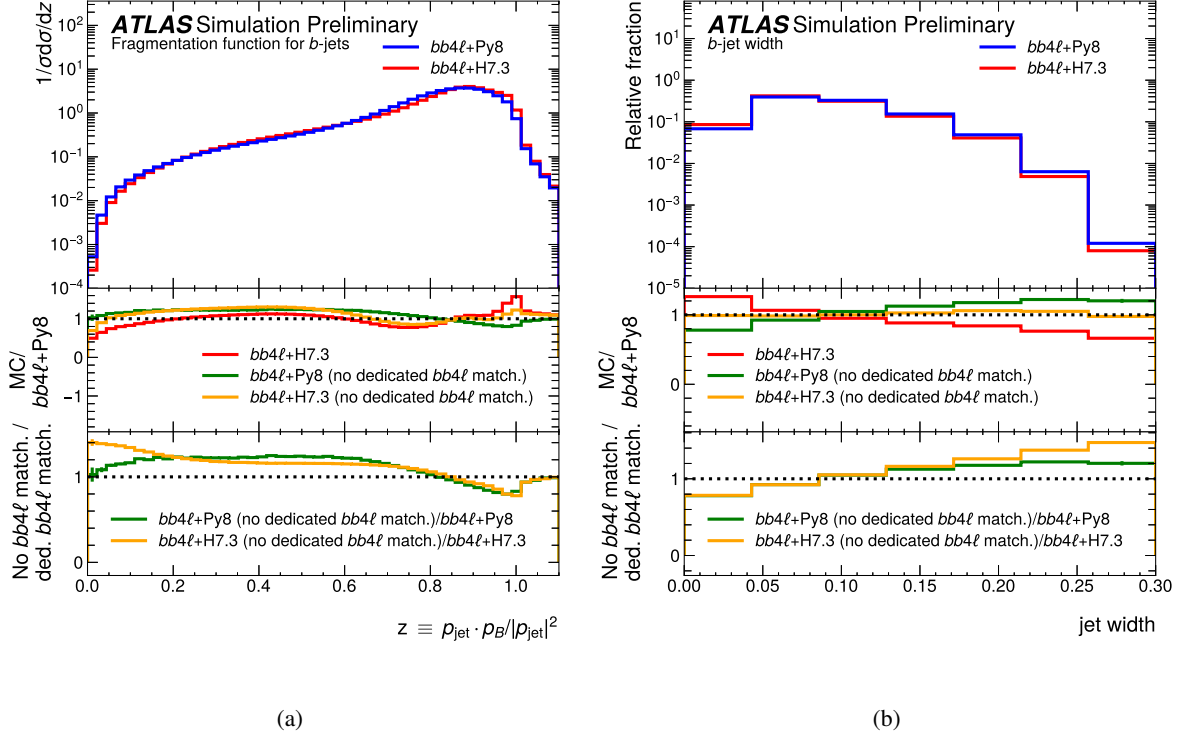


Figure 4: Comparison of the predictions for (a) the b -jet fragmentation function and for (b) the b -jet width of the nominal ATLAS $bb4\ell$ +PYTHIA 8 and the $bb4\ell$ +HERWIG 7 MC sample. Additionally, two MC samples in which $bb4\ell$ POWHEG LHE events are matched to either PYTHIA 8 or HERWIG 7 are produced, in which the required dedicated $bb4\ell$ matching is not applied, but instead the default POWHEG matching is used - which is applied e.g. in the $t\bar{t}$ hvq MC sample. A dedicated $bb4\ell$ parton shower matching is needed, since the allrad scheme was used in the $bb4\ell$ POWHEG event generation, which allows up to three emissions: POWHEG emissions from the production, the top-quark and the anti-top quark decay process. Consequently, three different matching scales need to be respected in the subsequent parton showering: While the emissions from the production process are vetoed via the default POWHEG matching procedure, the emissions from the top-quark and anti-top quark decay are vetoed via the dedicated parton shower matching. Thus, applying or not applying the dedicated $bb4\ell$ matching influences especially b -jet property distributions. Without the dedicated $bb4\ell$ matching, typically more parton shower emissions from the top-quark decay products are expected, and thus less b -hadrons with a large momentum fraction of the b -jet and broader b -jets are simulated without the dedicated matching. Further, the influence of the dedicated $bb4\ell$ matching is quite similar between the $bb4\ell$ predictions matched with PYTHIA and HERWIG, validating the two completely independent implementations of this $bb4\ell$ matching procedure.

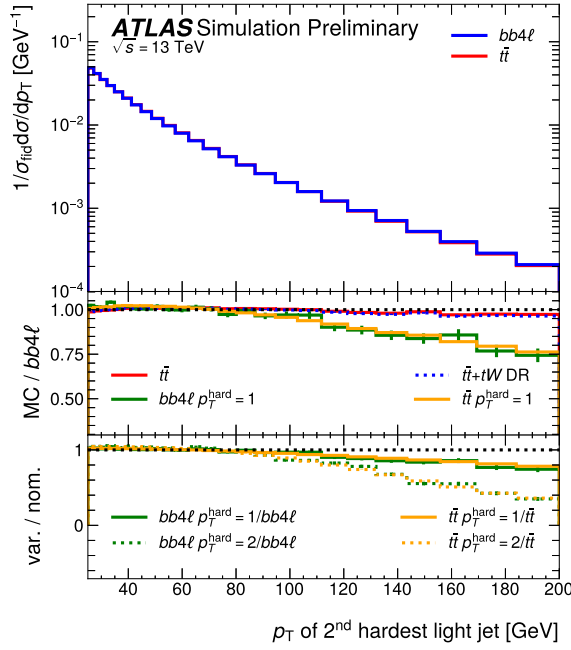


Figure 5: Comparison of the predictions for the transverse momentum (p_T) distribution of the second-hardest light jet of the nominal ATLAS $bb4\ell$, $t\bar{t}$ and $t\bar{t} + tW$ POWHEG+PYTHIA 8 sample, using either the default matching parameter $p_T^{\text{hard}} = 0$ or varying this parameter to $p_T^{\text{hard}} = 1$ and $p_T^{\text{hard}} = 2$ in the $bb4\ell$ and $t\bar{t}$ MC sample generation. The corresponding matching uncertainty is defined by comparing the nominal sample ($p_T^{\text{hard}} = 0$) to the varied MC sample with $p_T^{\text{hard}} = 1$. While the hardest additional jet in the event is expected to be mostly influenced by the hard process generation, the second hardest light jet should be sensitive to the parton showering and matching. Further, the hardest parton shower emissions are expected to originate from the production process, for which the same matching procedure is applied in $bb4\ell$ and $t\bar{t}$. This explains the excellent agreement in the uncertainty estimate in $bb4\ell$ and $t\bar{t}$ obtained by varying the p_T^{hard} parameter.

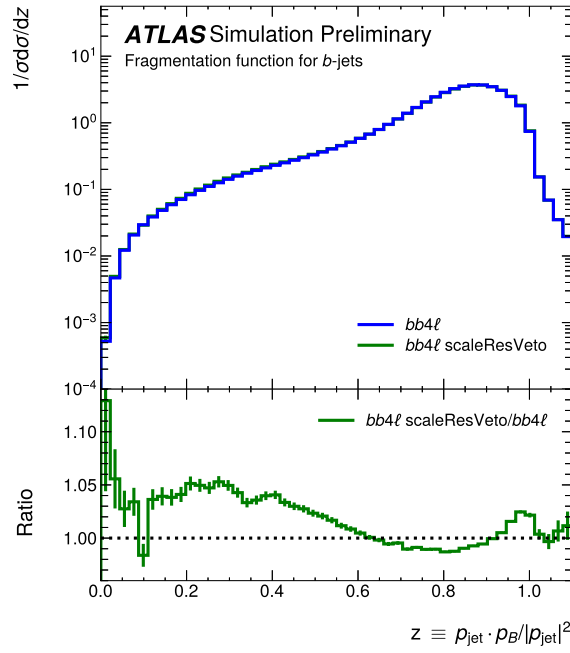


Figure 6: Comparison of the prediction for the b -jet fragmentation function of the nominal ATLAS $bb4\ell$ MC sample and a $bb4\ell$ MC sample, in which the matching in the top-quark decay is performed using the scale-resonance-veto option of the dedicated PYTHIA 8 UserHook, which implements the dedicated $bb4\ell$ matching for POWHEG events generated with the allrad scheme. Since this variation in the parton shower matching is only affecting emissions from the top-quark decay products, especially b -jet distributions should be sensitive to this variation. This is indeed found to be the case, shown here for the b -jet fragmentation function.

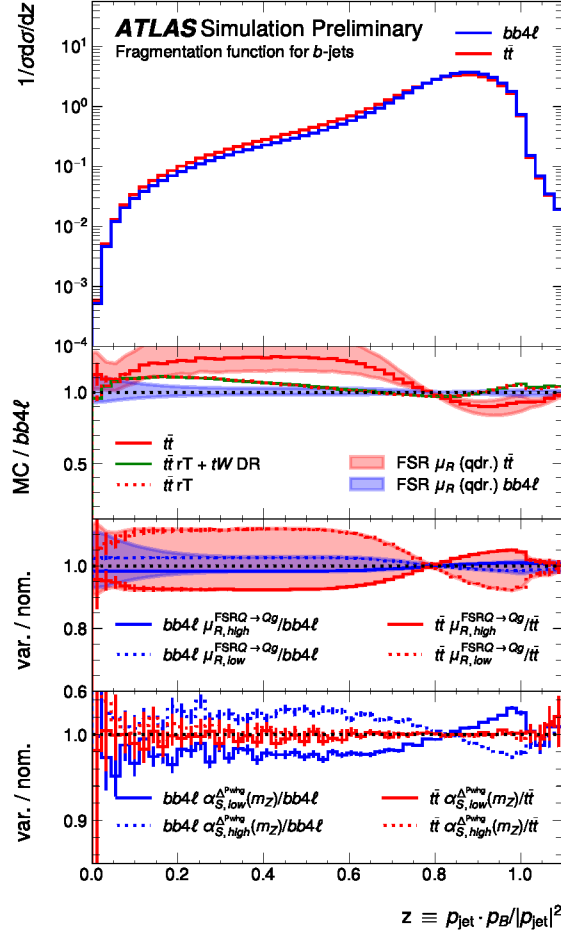


Figure 7: Comparison of the predictions of the b -jet fragmentation functions of the nominal ATLAS $bb4\ell$ sample, the nominal $t\bar{t}$ MC sample and the $t\bar{t}$ MC sample generated with the recoil-to-top (rT) PYTHIA 8 setting combined with the tW DR sample. The FSR uncertainty bands for $bb4\ell$ (blue) and $t\bar{t}$ (red) displayed in the ratio panels are obtained by varying μ_R in the separate splitting kernels ($g \rightarrow gg$, $g \rightarrow qq$, $q \rightarrow qg$, $Q \rightarrow Qg$ with $Q \in \{t, b\}$) and adding these four variations in quadrature. The full FSR μ_R uncertainty bands are compared to the variation of μ_R only in the FSR heavy quark splitting kernels ($Q \rightarrow Qg$) in the second ratio panel ($\mu_{R,\text{high}}^{\text{FSR},Q \rightarrow Qg} = 2\mu_0$ and $\mu_{R,\text{low}}^{\text{FSR},Q \rightarrow Qg} = 0.5\mu_0$). In the lowest ratio panel, the POWHEG Sudakov variation uncertainty is shown, evaluated by taking the difference between the nominal MC sample reweighted to the alternative PDF (with different $\alpha_S(m_Z) = 0.115$ (solid) and $\alpha_S(m_Z) = 0.121$ (dashed) value) to the MC sample explicitly generated with the alternative PDF set. The PDF reweighting does not affect the POWHEG Sudakov factor as discussed e.g. in [22], while explicitly using a PDF set with an alternative $\alpha_S(m_Z)$ value during event generation does. The differences in these variations between the $bb4\ell$ and $t\bar{t}$ predictions are interpreted as originating from the differences in the generation of the first emission from the top-quark decay: While the first emission from the top-quark decay products is always performed by the PYTHIA 8 parton shower in the $t\bar{t}$ MC sample, in case of the $bb4\ell$ simulation this emission can also be generated by POWHEG in the allrad emission generation scheme. Thus, in case of the $t\bar{t}$ MC sample, the FSR PYTHIA 8 μ_R $Q \rightarrow Qg$ splitting kernel variations influence the first emission from the top-quark decay - and thus from the b -parton from the top-quark decay - strongly. But for the $bb4\ell$ MC samples, this variation often only influences the second and subsequent emissions from top-quark decay, while the first emission is sensitive to variations of the POWHEG Sudakov factor. This interpretation is further supported by the fact the variations in the b -jet fragmentation function through the FSR μ_R $Q \rightarrow Qg$ splitting kernel variations and through the $\alpha_S(m_Z)$ variations in the POWHEG Sudakov are in good agreement based on the comparison of the shape of the ratios.

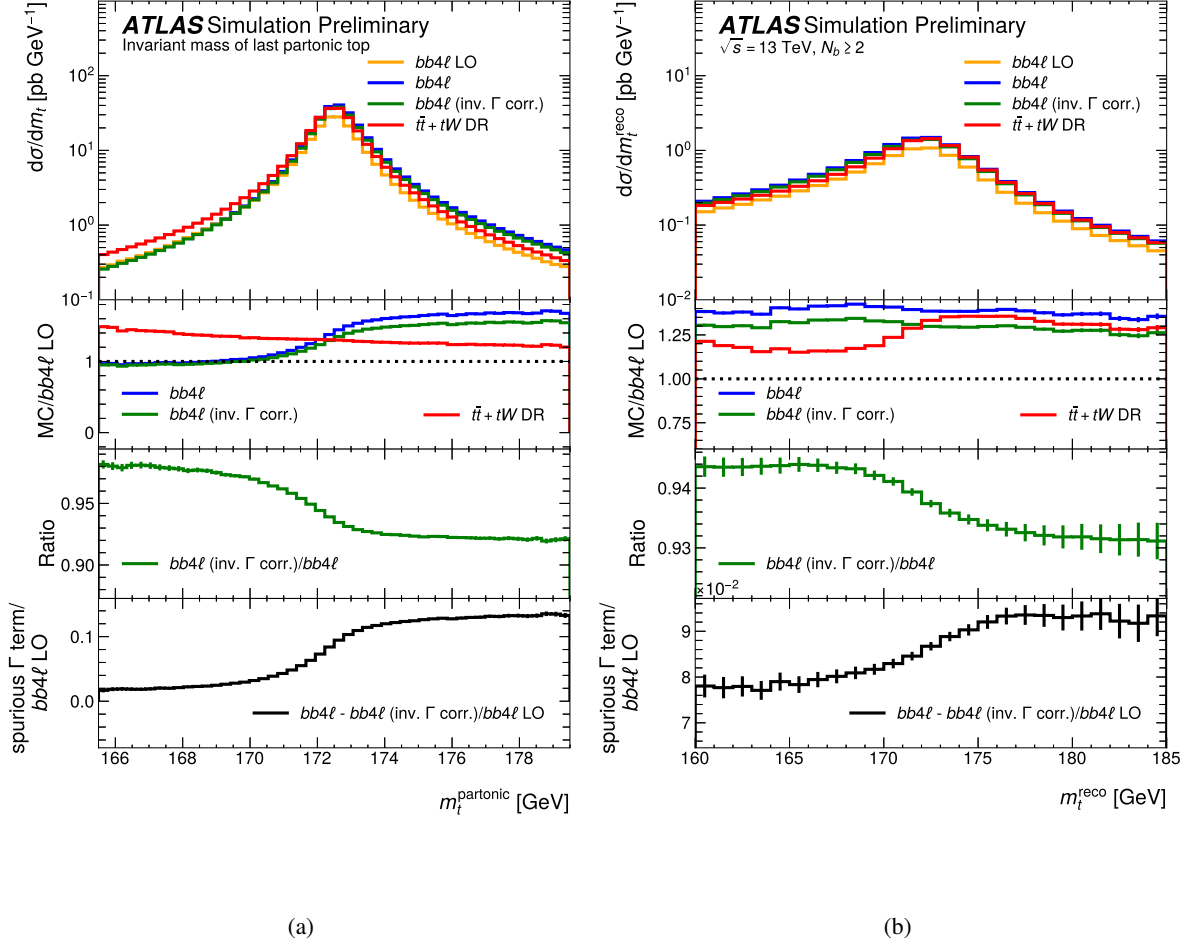


Figure 8: Comparison of the predictions of (a) the invariant mass distribution of the last parton-level top-quark in the event record and of (b) the invariant mass of the reconstructed top-quark of the nominal $bb4l$ and $t\bar{t} + tW$ ATLAS MC samples. Further, the comparison of these predictions to the ones obtained by applying either the inverse width correction weights in the $bb4l$ MC sample or generating a LO $bb4l$ MC sample, matched to the PYTHIA 8 parton shower is shown. Applying the inverse width corrections in the $bb4l$ simulation leads to a significant normalisation difference, which is in agreement with the findings in Ref. [2]. The bottom panel displays the spurious width effects, obtained as the difference between the nominal $bb4l$ prediction and the $bb4l$ predictions with the inverse width corrections, i.e. with the spurious width effects removed. To construct a similar term as in Eq. (3.16) in Ref. [2], this difference is divided by the $bb4l \text{ LO}$ prediction. It should be emphasized however, that in contrast to the predictions in Eq. (3.16) in Ref. [2], the $bb4l$ prediction does not correspond to a narrow-width approximation calculation. Especially in the invariant mass distribution of the parton-level top-quark, a pronounced shape variation is observed when applying the inverse width correction. The shape variation in the spurious width effects in the bottom panel does follow the shape of the differential $bb4l \text{ NLO/LO } K$ -factor. In Ref. [2] it was stated, that the spurious width effects are expected to follow the differential NLO/LO K -factor in narrow-width approximation calculations. This comparison of the spurious width terms and the NLO/LO K -factor in the parton-level top-quark invariant mass distribution is not affected by the different parton shower matching when interfacing NLO or LO $bb4l$ events with PYTHIA 8. Instead, this would affect the invariant mass of the reconstructed top-quark at particle level. Still, this distribution is interesting to study, as it is found, that the shape variation induced by the inverse width corrections is strongly reduced.

References

- [1] T. Ježo, J. M. Lindert, P. Nason, C. Oleari and S. Pozzorini, *An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects*, *Eur. Phys. J. C* **76** (2016) 691, arXiv: [1607.04538 \[hep-ph\]](#) (cit. on p. 2).
- [2] T. Ježo, J. M. Lindert and S. Pozzorini, *Resonance-aware NLOPS matching for off-shell $t\bar{t} + tW$ production with semileptonic decays*, *JHEP* **10** (2023) 008, arXiv: [2307.15653 \[hep-ph\]](#) (cit. on pp. 2–4, 7, 10, 15).
- [3] ATLAS Collaboration, *Observation of quantum entanglement with top quarks at the ATLAS detector*, *Nature* **633** (2024) 542, arXiv: [2311.07288 \[hep-ex\]](#) (cit. on p. 2).
- [4] ATLAS Collaboration, *Observation of a cross-section enhancement near the $t\bar{t}$ production threshold in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, ATLAS-CONF-2025-008, 2025, URL: <https://cds.cern.ch/record/2937636> (cit. on p. 2).
- [5] ATLAS Collaboration, *Measurements of differential cross-sections of $WbWb$ production in the dilepton channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, (2025), arXiv: [2506.14700 \[hep-ex\]](#) (cit. on p. 2).
- [6] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](#) (cit. on p. 2).
- [7] 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. 2).
- [8] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029, arXiv: [0805.3067 \[hep-ph\]](#) (cit. on p. 2).
- [9] ATLAS Collaboration, *Studies on the improvement of the matching uncertainty definition in top-quark processes simulated with POWHEG+PYTHIA8*, ATL-PHYS-PUB-2023-029, 2013, URL: <https://cds.cern.ch/record/2872787> (cit. on pp. 2, 5).
- [10] ATLAS Collaboration, *Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^+\ell'^-\nu\bar{\nu}'$ final states with POWHEG and MADGRAPH5_AMC@NLO setups*, ATL-PHYS-PUB-2021-042, 2021, URL: <https://cds.cern.ch/record/2792254> (cit. on pp. 2, 3).
- [11] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, *JHEP* **12** (2015) 065, arXiv: [1509.09071 \[hep-ph\]](#) (cit. on p. 2).
- [12] 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. 2).
- [13] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, arXiv: [0707.3088 \[hep-ph\]](#) (cit. on p. 2).
- [14] E. Re, *Single-top Wt -channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547, arXiv: [1009.2450 \[hep-ph\]](#) (cit. on pp. 2, 3).
- [15] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](#) (cit. on p. 3).

- [16] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, arXiv: [1410.3012 \[hep-ph\]](#) (cit. on p. 3).
- [17] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, *SciPost Phys. Codeb.* (2022) 8, arXiv: [2203.11601 \[hep-ph\]](#) (cit. on p. 3).
- [18] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419> (cit. on p. 3).
- [19] M. R. Masouminia and P. Richardson, *Hadronization and decay of excited heavy hadrons in Herwig 7*, *JHEP* **07** (2024) 278, arXiv: [2312.02757 \[hep-ph\]](#) (cit. on p. 3).
- [20] H. Brooks and P. Skands, *Coherent showers in decays of colored resonances*, *Phys. Rev. D* **100** (2019) 076006, arXiv: [1907.08980 \[hep-ph\]](#) (cit. on p. 4).
- [21] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 (cit. on p. 4).
- [22] S. Ferrario Ravasio, T. Ježo, P. Nason and C. Oleari, *A theoretical study of top-mass measurements at the LHC using NLO+PS generators of increasing accuracy*, *Eur. Phys. J. C* **78** (2018) 458, [Addendum: *Eur.Phys.J.C* 79, 859 (2019)] [A slightly modified version of the *bb4l*+Herwig7 matching source code was kindly provided to the ATLAS Collaboration by Silvia Ferrario Ravasio.], arXiv: [1906.09166 \[hep-ph\]](#) (cit. on pp. 4, 5, 14).
- [23] S. Mrenna and P. Skands, *Automated parton-shower variations in PYTHIA 8*, *Phys. Rev. D* **94** (2016) 074005, arXiv: [1605.08352 \[hep-ph\]](#) (cit. on p. 4).
- [24] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026, arXiv: [1912.05451 \[hep-ph\]](#) (cit. on p. 5).
- [25] Rivet, *MC_HFDECAYS*, https://rivet.hepforge.org/analyses/MC_HFDECAYS.html, Accessed: September,2024 (cit. on p. 5).
- [26] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#) (cit. on p. 5).
- [27] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, arXiv: [1111.6097 \[hep-ph\]](#) (cit. on p. 5).
- [28] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005, arXiv: [0802.1188 \[hep-ph\]](#) (cit. on p. 5).