ATLAS PUB Note

ATL-PHYS-PUB-2022-006

March 1, 2022

Studies of Monte Carlo predictions for the $t\bar{t}b\bar{b}$ process

The ATLAS Collaboration

Studies of Monte Carlo generator predictions are presented for top quark pair production in association with bottom quarks ($t\bar{t}b\bar{b}$). The calculations are performed at next-to-leading order in QCD for the $t\bar{t}b\bar{b}$ matrix-element using the four-flavour scheme. The parameters of the POWHEG+PYTHIA8 samples are optimised and the results are compared to predictions from POWHEG+HERWIG7 and SHERPA. The studies include comparisons to simulations at parton level with stable top quarks, to simulations at particle level and to unfolded data with the aim to implement the improved theoretical knowledge, yield better agreement with data and define systematic uncertainties for the modelling of the $t\bar{t}b\bar{b}$ process.

1 Introduction

Monte-Carlo (MC) predictions of top quark pair production in association with a bottom-quark pair ($t\bar{t}b\bar{b}$) suffer from large theoretical uncertainties and are a dominant uncertainty in measurements where this process is a significant background, such as top quark pair production in association with a Higgs Boson decaying to bottom quarks $t\bar{t}H(H \to b\bar{b})$ and four top quark production. The calculation of the $t\bar{t}b\bar{b}$ process is available at next-to-leading order (NLO) QCD [1, 2] and the implementation in the POWHEG BOX RES framework [3] (referred to as POWHEG in the following) has been used in ATLAS physics analyses [4, 5]. The simulation of inclusive top-quark pair production using POWHEG and the detailed parameter settings were studied in [6] for this process. However, due to the multi-scale nature of the $t\bar{t}b\bar{b}$ process, special parameter settings are needed. The study presented here optimises the parameter settings in the POWHEG $t\bar{t}b\bar{b}$ calculation and its matching to the PYTHIA8 [7] and HERWIG7 [8] parton shower (PS) based on theoretical arguments and comparison to data when appropriate and possible. Furthermore, comparisons are presented to the corresponding prediction of SHERPA 2.2.10 [9–11]. The aim is to yield better agreement of the predictions with data at a center-of-mass energy of 13 TeV and develop a scheme of modelling uncertainties for the next round of data analyses.

The comparisons are performed at particle level. In the parton level case, special samples where the top quarks[1] are treated as stable. This treatment is avoiding the combinatorial complexity of identifying jets initiated by b-, charm or light quarks from the top quark decay chain and of separating them from additional jets not originating from the decay of a top quark. These additional jets are expected to be the primary cause of the differences in modelling. All comparisons are done using the Rivet analysis toolkit [12].

The note is structured as follows: Section 2 will list the generator settings for the samples with decayed and with stable top quarks. Section 3 gives an overview of the object definitions and the kinematic phase space analysed. The settings related to the calculations in POWHEG such as the scale choice, the damping and the integration parameters as well as the settings of the top quark decays are studied in Section 4, while Section 5 describes studies related to the settings in the matched parton shower. Section 6 shows comparisons of the generators considered to estimate modelling uncertainties of this process in typical analysis phase spaces. Section 7 will summarize the findings of this note and compare the different $t\bar{t}b\bar{b}$ predictions to estimate modelling uncertainties on observables relevant for the $t\bar{t}b\bar{b}$ and $t\bar{t}H(H \to b\bar{b})$ measurements.

2 Generators and their settings

2.1 Common generator settings

For the generation of the $t\bar{t}b\bar{b}$ process, b-quarks are treated as massive with a mass of $m_{\mathrm{b}} = 4.75\,\mathrm{GeV}$ and correspondingly, a set of parton distribution function (PDF) based on the four flavour scheme is used. The top quark mass is set to $m_{\mathrm{top}} = 172.5\,\mathrm{GeV}$. The matrix element (ME) calculation in POWHEG is either matched to PYTHIA8 with the A14 set of tuned parameters [13] or HERWIG7 for the PS, multiparton-interaction, beam remnant and hadronisation processes. Unless noted otherwise, top quark decays were simulated at LO using the code provided in POWHEG to preserve all spin correlations. B-hadron decays are

[1] "top quarks" refers to both top- and antitop quarks

simulated using EvtGen [14]. The same set of POWHEG LHE files are used for matching to the varied shower settings inside PYTHIA8 and for matching to HERWIG7.

Finally comparisons to SHERPA 2.2.10 [10] which uses the same OPENLOOPS [11] code for the ME calculation and the same scale settings but implements a MC@NLO [15] type matching and its own PS and particle decays are made. The MC generators used in this note and their settings are summarised in Table 1.

Table 1: Configurations used for the MC generation.

Generator	ME order	PS	PDF	Tune
POWHEG BOX RES	NLO	PYTHIA 8.244	NNPDF3.1 nnlo Nf4	A14
	NLO	HERWIG 7.1.6	NNPDF3.1 nnlo Nf4	H7.1.6 default
SHERPA 2.2.10	NLO	SHERPA	NNPDF3.0 nnlo Nf4	SHERPA default

2.2 Simulating predictions with stable top quarks

In addition to the samples mentioned above special samples were simulated with the POWHEG+PYTHIA8 settings described above but where the top quarks were not decayed. The POWHEG setup and the settings remain unchanged. However, since the top quarks are not decayed, hadronisation is switched off as well. Furthermore, multi-parton interaction (MPI) and QED radiation in the PS are switched off following the settings in Ref. [3].

3 Kinematic regions and observables

Comparisons to ATLAS measurements of the $t\bar{t}b\bar{b}$ cross sections are used to discriminate and exclude certain generator set-ups wherever possible. For the MC only comparisons, the studied observables and kinematic regions are selected with the dual focus to be sensitive to the details of the MC models and to match the analysis phase space used for $t\bar{t}b\bar{b}$ [16] and $t\bar{t}H(H \to b\bar{b})$ [4] measurements.

Every analysis follows the commonly used ATLAS object definition at particle level, i.e. all objects are defined using stable final state particles with a lifetime of at least 30 ps. Jets are reconstructed from all stable final state particles (but excluding leptons and neutrinos from the top quark decay chain) using the anti-k_t jet algorithm [17] with a radius parameter of $R = 0.4$. Jets which contain at least one ghost-associated [18] B-hadron with transverse momentum (p_T) of $p_T \geq 5$ GeV are defined as b-jets. The four-momentum of the bare leptons from top quark decay are modified ("dressed") by adding the four-momenta of all radiated photons within a cone of size $\Delta R = 0.1$. All objects are considered within pseudo-rapidity $|\eta| \leq 2.5$ and with $p_T > 27$ GeV for leptons and $p_T > 25$ GeV for jets and b-jets .

Since the samples with stable top quarks were generated with hadronisation switched off, jets are built from the final state partons using the anti-k_t jet algorithm with a radius of 0.4. b-jets are defined as jets which contain at least one ghost-matched b-quark with $p_T > 5$ GeV. The same kinematic requirements are applied on these objects as for decayed top quarks and are listed in Table 2.

Regions of phase space were defined requiring a high number of jets and b-jets corresponding to the particles produced in the Born level process. These regions are similar to those used in the $t\bar{t}b\bar{b}$ and

Table 2: Kinematic requirement on the objects of the analyses.

Defined object	Kinematic restrictions
Leptons (electrons and muons)	$p_{\mathrm{T}} > 27$ GeV and $\lvert\eta\rvert < 2.5$
Jets (including b-jets)	$p_{\mathrm{T}} > 25$ GeV and $\lvert\eta\rvert < 2.5$

$t\bar{t}H(H \to b\bar{b})$ measurements and include the $t\bar{t}H(H \to b\bar{b})$ signal regions. The regions are defined in the following sections.

In case of stable top quarks, in addition to the regions with one and two b-jets regions with additional light jets are defined since these regions are expected to be especially sensitive to modelling effects. A similar set of observables is studied for each region. Unless otherwise specified jets refers to any flavour jet including light- and b-jets.

3.1 $t\bar{t}b\bar{b}$ analysis with decayed top quarks

This analysis uses $t\bar{t}b\bar{b}$ events at particle level with decayed top quarks. This analysis was developed to compare ATLAS and CMS $t\bar{t}b\bar{b}$ predictions in the context of the LHC Higgs working group (WG). The objects stated in Table 2 are used to define different kinematic regions in which observables are studied. The regions are separated in the single lepton and the dilepton $t\bar{t}$ decay channels and according to the b-jet multiplicities. The selection with exactly three b-jets has a higher contribution of events where two b-hadrons are merged into one jet, compared to the region with four b-jets.

The regions are:

1. 1l 3b 5j: 1 lepton (e/μ), 3 b-jets $\geq$ 5 jets.
2. 1l 4b 6j: 1 lepton (e/μ), $\geq$ 4 b-jets $\geq$ 6 jets.
3. 2l 4b 4j: 2 leptons of opposite electric charge (e/μ), $\geq$ 4 b-jets $\geq$ 4 jets.

The observables used in each region are summarized in Table 3.

Table 3: Descriptions of the observables used for the analysis with decayed top quarks.

Observable	Description
Number of jets	Number of jets inclusive in jet flavour
Number of b-jets	Number of b-jets
Leading b-jet p_{T}	Transverse momentum of leading b-jet
Leading light jet p_{T}	Transverse momentum of the leading light jet
Subleading light jet p_{T}	Transverse momentum of the subleading light jet
H_{T}	Scalar sum of all transverse momenta over leptons and jets
ΔR_{bb}	$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ for the two b-jets with the smallest angular difference
$\eta_{\mathrm{jj}}^{\mathrm{max}}$	Highest $\Delta\eta$ value between any combination of two jets
m_{bb}	Mass of the two b-jets with the smallest angular difference
$\Delta\phi(l^+, l^-)/\pi$	Angular difference between the two leptons of opposite electric charge

Since at least three b-jets are required in all regions, two b-jets have to be selected for the calculation of ΔR_{bb} and m_{bb} . At particle level, the two b-jets which yield the smallest ΔR value are taken.

3.2 $t\bar{t}b\bar{b}$ analysis with stable top quarks

This analysis studies $t\bar{t}b\bar{b}$ events where the top quarks are not decayed as described in Section 2. In the studies of $t\bar{t}b\bar{b}$ events with stable top quarks the same set of kinematic requirements as for the jets are imposed on the top quarks. Kinematic regions are defined dependent on the number of b-jets and light jets:

1. ttb : $t\bar{t}$ + $\geq 1b$-jet
2. ttbb: $t\bar{t}$ + $\geq 2b$-jets
3. ttbj: $t\bar{t}$ + $\geq 1b$-jet + $\geq$ 1 light jet
4. ttbbj: $t\bar{t}$ + $\geq 2b$-jets + $\geq$ 1 light jet

For observables like ΔR_{bb} and m_{bb} at least two b-jets are required. This converts the ttb region into an effective ttbb region for these observables. Similar observables to the analysis with decayed top quarks are studied as listed in Table 4.

Table 4: Descriptions of the observables used for the analysis with stable top quarks.

Observable name	Descriptions of the observables
Number of light jets	Number of light jets
Number of b-jets	Number of b-jets
Leading b-jet p_{T}	Transverse momentum of the leading b-jet
Subleading b-jet p_{T}	Transverse momentum of the subleading b-jet
Leading light jet p_{T}	Transverse momentum of the leading light jet
Subleading light jet p_{T}	Transverse momentum of the subleading light jet
H_{T}	Scalar sum of all jet transverse momenta (top quarks are not included in the sum)
ΔR_{bb}	$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ for the two b-jets with the smallest angular difference
m_{bb}	Mass of the two b-jets with the smallest angular difference
$\eta_{\mathrm{jj}}^{\mathrm{max}}$	Highest η value between any combination of two jets

Observables that are requiring two b-jets are not calculated in the regions with only one b-jet

3.3 Comparison to data

The MC predictions are compared to the ATLAS measurements of the fiducial and differential cross-section of $t\bar{t}b\bar{b}$ production in proton-proton collisions with a centre-of-mass energy of 13 TeV in the semileptonic and dileptonic channel using 36.1 fb^{-1}[16]. The objects in this measurement are reconstructed as described above with the kinematic requirements as listed in Table 2. The following fiducial regions are defined

1. ℓ+jets $\geq$ 3b: 1 lepton (e/μ), $\geq$ 3 b-jets $\geq$ 5 jets.
2. ℓ+jets $\geq$ 4b: 1 lepton (e/μ), $\geq$ 4 b-jets $\geq$ 6 jets.

3. e$\mu \geq$ 3b: 2 leptons of opposite electric charge (e/μ), $\geq$ 3 b-jets $\geq$ 3 jets.

4. e$\mu \geq$ 4b: 2 leptons of opposite electric charge (e/μ), $\geq$ 4 b-jets $\geq$ 4 jets.

The observables which are unfolded to particle level are listed in Table 5.

Table 5: Descriptions of the observables used for the ATLAS $t\bar{t}b\bar{b}$ analysis

Observable name	Descriptions of the observables
σ_{fid}	Fiducial cross-section
H_{T}	Scalar sum of all transverse momenta over leptons and jets
Leading b-jet p_{T}	Transverse momentum of the leading b-jet.
ΔR_{bb}	$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ for the two b-jets with the smallest angular difference

4 Optimising POWHEG settings

The code provided in the POWHEG generator provides a calculation of the $t\bar{t}b\bar{b}$ process at NLO in QCD based on the POWHEG method which implements NLO corrections to the hardest emission to overcome the problem of negative weighted events as observed in other NLO predictions [19]. The calculation involves several free parameters and model choices which are studied and presented in the following. Some of these parameters have been optimised for inclusive $t\bar{t}$ production, but might differ for $t\bar{t}b\bar{b}$ due to the specifics of this multi-scale process.

This section is structured as follows: in Section 4.1 scale settings for the NLO ME calculation are studied. Section 4.2 discusses the free parameters of the damping function used in the POWHEG method to separate the singular from the finite parts of the real-emission matrix elements. Section 4.3 compares the default POWHEG implementation of the top quark decays against an alternative implementation provided by MadSpin [20]. Finally, Section 4.4 reports on the optimised POWHEG integration parameters to obtain a good compromise between run time and the fraction of negative-weighted events.

Table 6 gives an overview of the studied parameters, the values used in the previous samples for the ATLAS measurements [4, 5] and the new default values that are obtained following the studies presented in the following sections 4.1 and 4.2. The PDF set was changed to a newer version as shown in Table 6.

4.1 Factorisation and renormalisation scale for $t\bar{t}b\bar{b}$ predictions

Since the $t\bar{t}b\bar{b}$ process scales with α_{S}^4 at LO, the $t\bar{t}b\bar{b}$ cross-section is highly sensitive to the choice of the renormalisation scale μ_{R}. The process involves two widely separated scales, the top quark pair production at around 350 GeV and bottom pair production at around 10 GeV. In order to guarantee that the strong coupling factors associated to the final state objects adapt to the respective transverse energies, the geometric average of the two scales was suggested as functional form of μ_{R} in Ref. [22] and is listed in Tab.6. The factorisation scale μ_{F} which defines the available phase space for QCD radiation, should be related to the average top quark transverse energy as suggested based on theoretical arguments in Ref. [23]. These choices for μ_{R} and μ_{F} are the default in Ref. [3] and have been used as nominal scale in Ref. [4, 21]. They are listed as "previous value" in Table 6.

Table 6: POWHEG parameters used in the previous and the new default following the studies presented in the following sections 4.1 and 4.2. The parameters for POWHEG are used for both POWHEG+PYTHIA8 and POWHEG+HERWIG7 predictions, the PYTHIA8 parameters only affect the samples matched to PYTHIA8. Previous value refers to the settings in the samples used for the ATLAS measurements in Ref. [4, 21]. For explanation of the parameters see text in the corresponding section.

Parameter	previous value	new default value
PDF	NNPDF3.0 nnlo Nf4	NNPDF3.1 nnlo Nf4
Scale choice[a] [b]	$\mu_{\mathrm{R}}^{\mathrm{def}} = \sqrt[4]{\prod_{i=t,\bar{t},b,\bar{b}} E_{\mathrm{T,i}}}$ $\mu_{\mathrm{F}}^{\mathrm{def}} = \frac{1}{2}[\sum_{i=t,\bar{t},b,\bar{b},j} E_{\mathrm{T,i}}]$	$\mu_{\mathrm{R}} = 0.5 \cdot \mu_{\mathrm{R}}^{\mathrm{def}}$ $\mu_{\mathrm{F}} = \mu_{\mathrm{F}}^{\mathrm{def}}$
h_{bzd}	2	5
h_{damp}	$H_{\mathrm{T}}/2$	$H_{\mathrm{T}}/2$
Decay handling	MadSpin	POWHEG
PYTHIA8 POWHEG:$pTdef$	2	1
PYTHIA8 SpaceShower:dipoleRecoil	off (global recoil)	off (global recoil)

[a] The transverse energy E_{T} is defined as $E_{\mathrm{T}} = \sqrt{m^2 + p_{\mathrm{T}}^2}$

[b] In the definition j refers to either a quark or a gluon.

However, analyses using predictions produced with these scale choices needed large corrections factors for the $t\bar{t}b\bar{b}$ background normalisation [4] indicating that they underestimate the data. Therefore we study the following settings for the nominal scale and compare them to unfolded $t\bar{t}b\bar{b}$ measurements with the aim to obtain an optimal description of the data:

- $scale1$: $\mu_{\mathrm{R}} = \mu_{\mathrm{R}}^{\mathrm{def}}$, $\mu_{\mathrm{F}} = \mu_{\mathrm{F}}^{\mathrm{def}}$ ("previous value", default in Ref. [3])
- $scale2$: $\mu_{\mathrm{R}} = 0.5 \cdot \mu_{\mathrm{R}}^{\mathrm{def}}$, $\mu_{\mathrm{F}} = \mu_{\mathrm{F}}^{\mathrm{def}}$ (new LHC Higgs WG recommendation)
- $scale3$: $\mu_{\mathrm{R}} = 0.5 \cdot \mu_{\mathrm{R}}^{\mathrm{def}}$, $\mu_{\mathrm{F}} = 0.5 \cdot \mu_{\mathrm{F}}^{\mathrm{def}}$

Figure 1 shows the predictions with the different scale choices compared to the ATLAS measurement of the $t\bar{t}b\bar{b}$ fiducial cross-section. The predictions using "scale 1" underestimate the measured cross-section by 25-40 %, similar to the observed scaling factors in the $t\bar{t}H(H \to b\bar{b})$ and $t\bar{t}t\bar{t}$ analysis. The "scale 2" and "scale 3" settings are describing the measured data within the rather large experimental uncertainties.

Recently a fixed order calculation of $t\bar{t}b\bar{b}j$ at NLO [24], i.e. one order higher in α_{s}, became available. Comparisons of the POWHEG predictions to this calculation were performed in the context of the LHC Higgs WG [25] and demonstrated that with the "scale 2" set-up the convergence of the calculation is improved as evidenced by the reduction of the NLO K-factor from 1.62 (for $t\bar{t}b\bar{b}j$) to 1.54 (for $t\bar{t}bj$) comparing the $t\bar{t}$ production cross-sections with at least two additional b-jet or at least one at NLO to LO. Furthermore, Ref. [25] shows that when using the scale 2 settings, the p_{T} spectrum of additional light jets and the rate of events with an additional light jet in the $t\bar{t}b\bar{b}$ POWHEG MC prediction agree better with the fixed order prediction. The spread in the distribution of the p_{T} of the additional light jets predicted by different MC codes in the study is also significantly reduced. The LHC Higgs WG therefore updated their recommendation to use the "scale 2" setting as the new default.

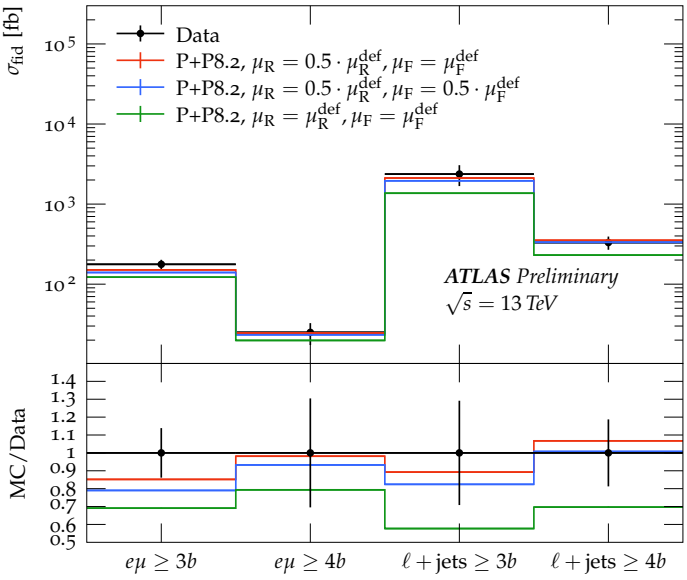

Figure 1: Fiducial cross-section of $t\bar{t}b\bar{b}$ as measured by ATLAS [16] compared to POWHEG+PYTHIA8 predictions with different scale settings as defined in the text. In this legend as well as all the legends of the following plots POWHEG+PYTHIA8 will be shortened to P+P8.2.

Since the data could not discriminate between "scale 2" and "scale 3" and to be consistent with the LHC Higgs working group recommendation, "scale 2" was chosen as the nominal scale choice for the new POWHEG $t\bar{t}b\bar{b}$ set-up. The SHERPA 2.2.10 sample was generated using the same scale definitions.

4.2 Steering the POWHEG damping function using h_{damp} and h_{bzd}

A key element of the POWHEG method is that the real emission part of the NLO calculation, R, is split into a finite part R_f corresponding to the fixed-order calculation of the resolved final state and a singular part R_s, containing the areas of the phase space where divergences may happen. The singular part resums large logarithms of soft and collinear emissions in a Sudakov form factor according to the POWHEG formula. A brief description of the parameters steering the transition between the two parts following closely [3] is given below.

The transition between the singular and the finite region is based on the p_T of the emitted parton and is regulated via a damping function[2] F which is defined to take on values between 0 and 1 such that

$$R_s = F \cdot R$$
$$R_f = [1 - F]R,$$

[2] F is defined in the full phase space for real emission, ϕ_R, i.e. $F = F(\phi_R)$

i.e. F effectively assigns events to the singular and the finite part of the calculation for $F \to 1$ and $F \to 0$ respectively. F is implemented as a product of two functions $F = F_{\text{damp}} \cdot F_{\text{bzd}}$, where F_{damp} and F_{bzd} address different types of divergences, see Ref. [26] for more details.

F_{damp} is defined as

$$F_{\text{damp}} = \frac{h_{\text{damp}}^2}{h_{\text{damp}}^2 + p_{\text{T}}^2}.$$

The parameter h_{damp} is a free model parameter that can be set by the user and p_{T} corresponds to the transverse momentum of the hardest parton from the real emission contribution. Soft and collinear emissions will lead to $F_{\text{damp}} \approx 1$ while hard emissions will be attributed to the finite part with $F_{\text{damp}} \to 0$ for $p_{\text{T}} \to \infty$. F_{damp} will smoothly change the weight of real radiation from R_s to R_f when the hardness of the emission (p_{T}) becomes of the order of h_{damp} or higher. The value of h_{damp} has been optimised by ATLAS to be 1.5 times the top quark mass (m_{top}) for $t\bar{t}$ inclusive production [27].

For the $t\bar{t}b\bar{b}$ process h_{damp} is set to $H_{\text{T}}/2$ following the recommendations in Ref [3]. Variations of the parameter between the values of H_{T}, $H_{\text{T}}/4$ and $1.5m_{\text{top}}$ are probed to estimate modelling uncertainties. We choose the typical factor 2 variations to estimate the scale uncertainties as suggested in Ref [3]. In addition, we included $1.5m_{\text{top}}$ as it is the choice used for the $t\bar{t}$ samples.

Figure 2 shows the effect of h_{damp} variations on the predictions when using the analysis with stable top quarks (see Section 3.2). The largest differences are observed when the value of h_{damp} is reduced by a factor of 2 to $H_{\text{T}}/4$ which yields an about 10% higher cross-section with respect to the $H_{\text{T}}/2$ prediction. This difference is however almost flat in the spectra of the studied observables. The difference between the $H_{\text{T}}/2$ and $H_{\text{T}}/4$ predictions is further reduced in the phase space analysed at particle level as demonstrated in Figure 3. In general Figure 3 shows a better agreement between the different h_{damp} variations.

In addition to the commonly known h_{damp} factor, a second function defined as

$$F_{\text{bzd}} = \theta\left(\mathrm{h}_{\text{bzd}} - \frac{R}{\mathcal{R}}\right)$$

is implemented where $\mathcal{R}$ corresponds to the infrared (soft and collinear) approximation of the full matrix element. F_{bzd} causes that in the vicinity of infrared singularities, where $R/\mathcal{R} \to 1$, radiative contributions are attributed to R_s and resumed. However, when the real emission matrix element exceeds the infrared approximation by a large factor of typically 2-10, the corresponding events are attributed to the finite remnant through the theta function as the resummation is not theoretically justified. Originally, this parameter was introduced to protect against non-physical divergences when the underlying Born cross-section approaches zero. However, similar enhancements can also happen in multi-scale processes such as $t\bar{t}b\bar{b}$ due to soft and collinear enhancements in $t\bar{t}b\bar{b}g$ events. In this process enhancements can arise in events with additional light-jet radiation, due to generation of a $b\bar{b}$ system at scales well below the hard energy of the full process, *i.e.* in regions where $m_{\text{bb}} \ll m_{\text{ttbb}}$ and/or $p_{\text{T,bb}} \ll m_{\text{ttbb}}$. The theta function causes that such events are shifted from the singular region into the finite region to avoid these non-physical enhancements and the amount of events that are shifted to the finite region can be steered via the parameter h_{bzd} which can be set by the user. Lower values of h_{bzd} will lead to more events in the finite region.

The dependence of the $t\bar{t}b\bar{b}$ kinematics on the particular choice of the h_{bzd} parameter value was investigated by varying h_{bzd} between 2, 5 and 10. Theoretical studies described in Ref [3] describe that high values of

h_{bzd} lead to a large scale dependence of the results. Therefore, a value of 2 was suggested by the authors in the original study [3]. However, after further studies and lowering the default scale as described in Section 4.1, the authors are now suggesting a h_{bzd} value of 5 as default as this leads only to a small increase of the scale dependence [28]. Variations of this parameter are considered as uncertainties as discussed in Ref. [3].

The dependence of the results on the particular choice of the parameter value is investigated by varying h_{bzd} for events with stable and decayed top quarks. In all figure the ratio is given in respect to the POWHEG+PYTHIA8 prediction. Figure 4 shows that the h_{bzd} variations lead to up to 15 % differences in some of the observables in the analysis with stable top quarks. As shown in Figure 5 these effects are reduced to about 10 % in the phase space of the analysis with decayed top quarks. The difference between $h_{\text{bzd}} = 5$ and $h_{\text{bzd}} = 10$ is largely reduced while the difference between $h_{\text{bzd}} = 5$ and $h_{\text{bzd}} = 2$ remains noticeable. The authors are suggesting a h_{bzd} value of 5, as also used for inclusive $t\bar{t}$ production. Therefore, the default will be changed to the newly recommended $h_{\text{bzd}} = 5$ value and $h_{\text{bzd}} = 2$ is considered as a modelling uncertainty.

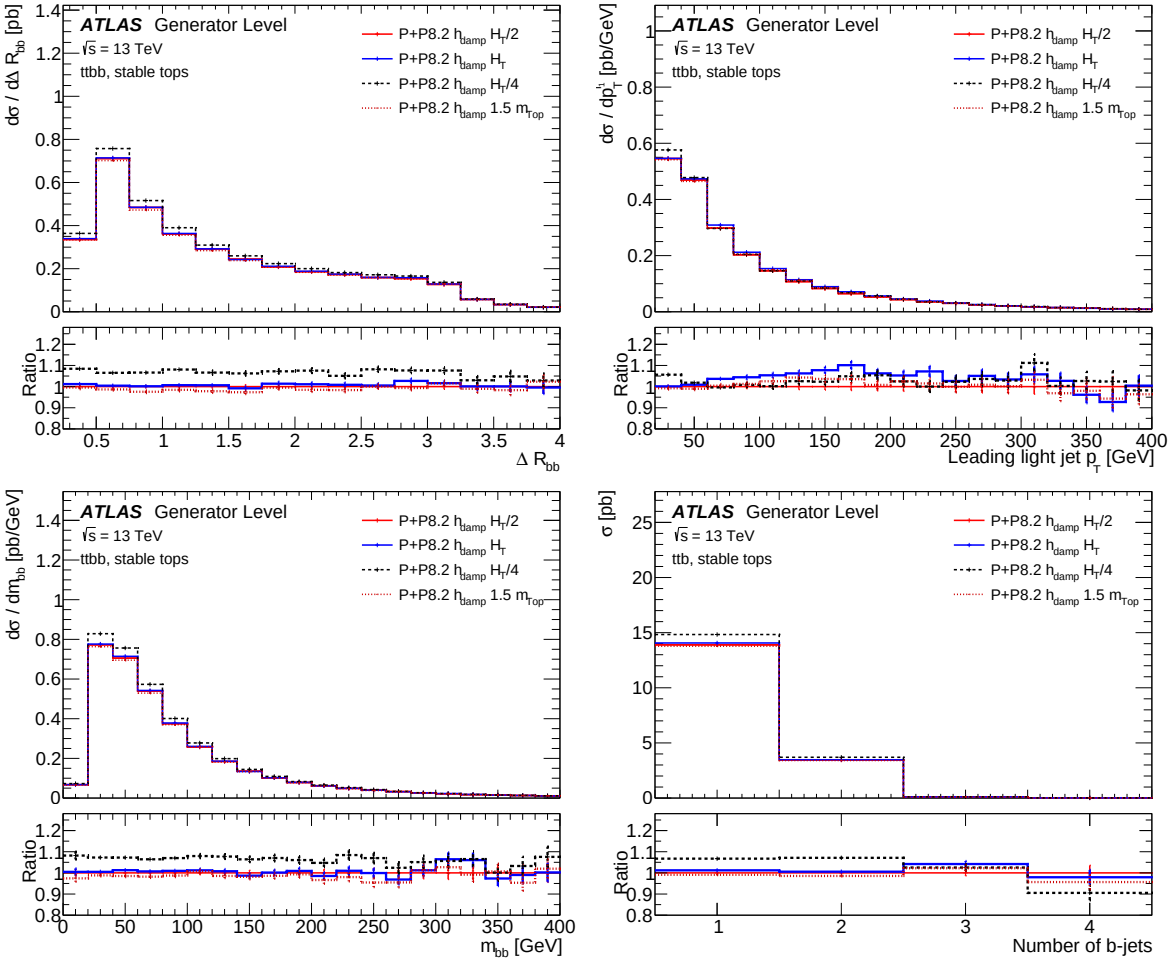

Figure 2: Kinematic distributions with different h_{damp} settings using predictions with stable top quarks in the $t\bar{t}b\bar{b}$ region as defined in Section 3.2. The ratio is computed with respect to the P+P8.2 $h_{\text{damp}} = H_{\text{T}}/2$ prediction. ΔR_{bb} between two leading b-jets (upper left), light jet p_{T} (upper right), invariant mass distribution between the two leading b-jets (lower left), b-jet multiplicity (lower right).

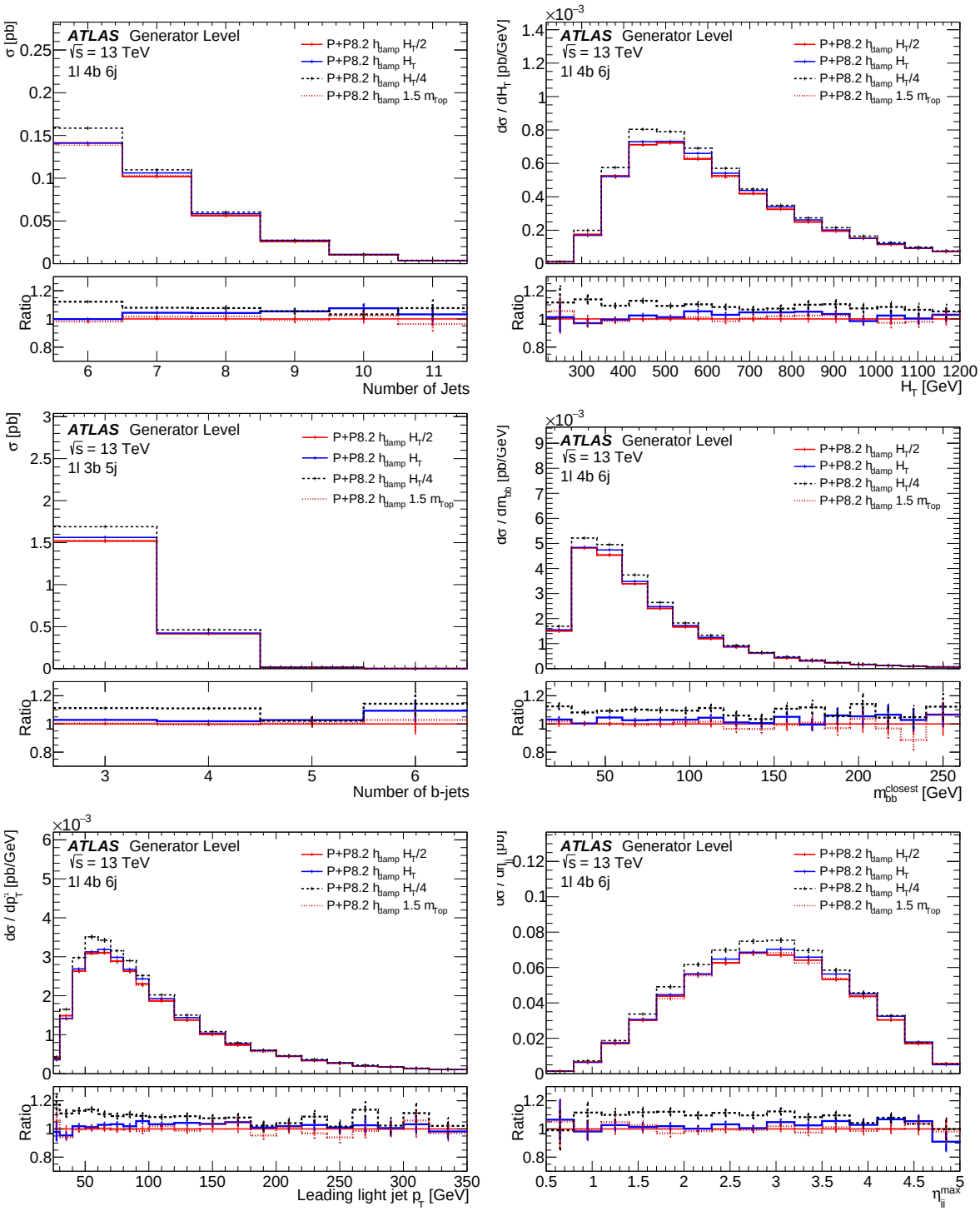

Figure 3: Kinematic distributions with different h_{damp} settings using predictions with decayed top quarks in the 1l 4b 6j region as defined in Section 3.1. The ratio is computed with respect to the P+P8.2 $h_{\text{damp}} = H_{\text{T}}/2$ prediction. Jet multiplicity (upper left), H_{T} (upper right), b-jet multiplicity (middle left), m_{bb} (middle right), light jet p_{T} (lower left), $\eta_{\text{jj}}^{\text{max}}$ (lower right).

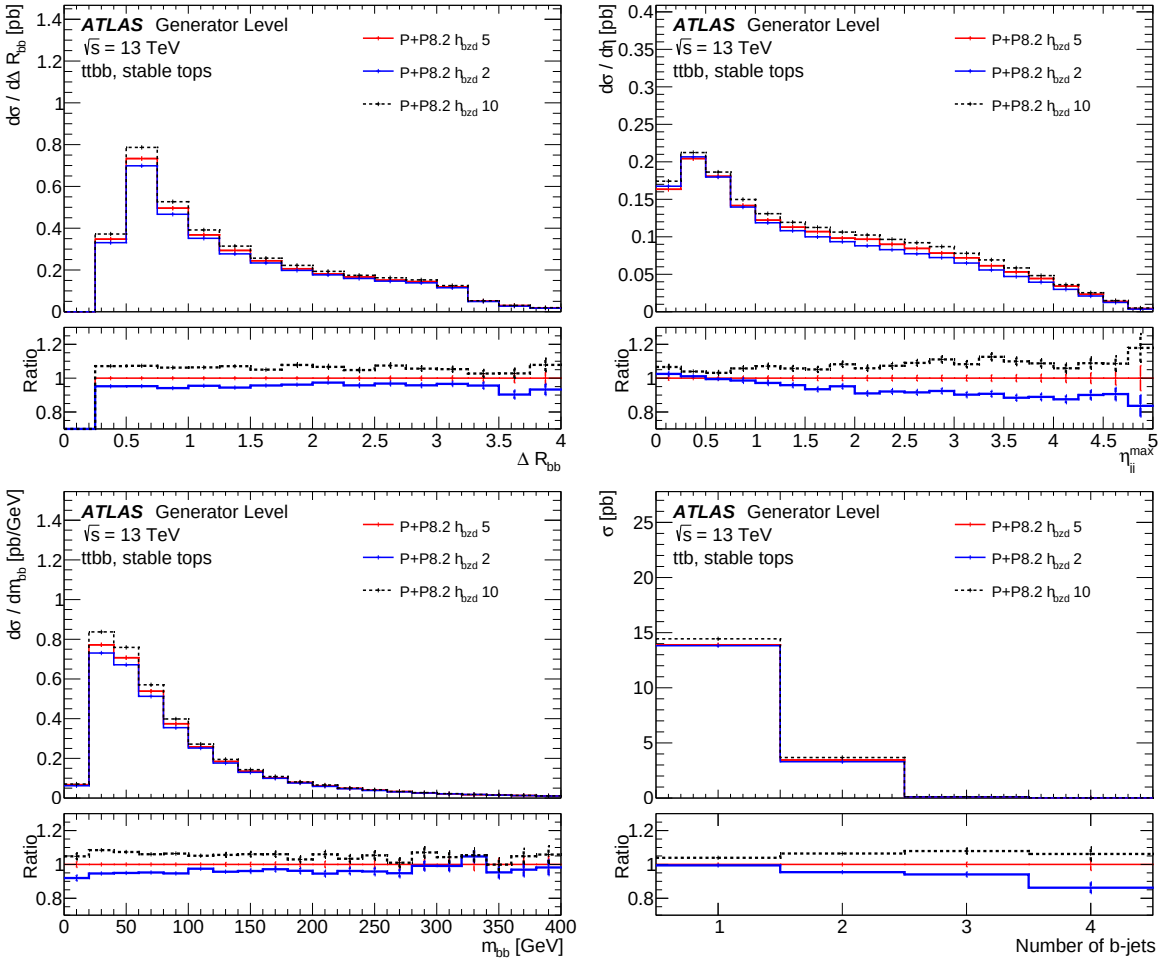

Figure 4: Kinematic distributions with different h_{bzd} settings using predictions with stable top quarks in the $t\bar{t}b\bar{b}$ region as defined in Section 3.2. The ratio is computed with respect to the P+P8.2 $h_{\mathrm{bzd}} = 5$ prediction. ΔR_{bb} between two leading b-jets (upper left), $\eta_{\mathrm{jj}}^{\mathrm{max}}$ (upper right), invariant mass distribution between the two leading b-jets (lower left), b-jet multiplicity (lower right).

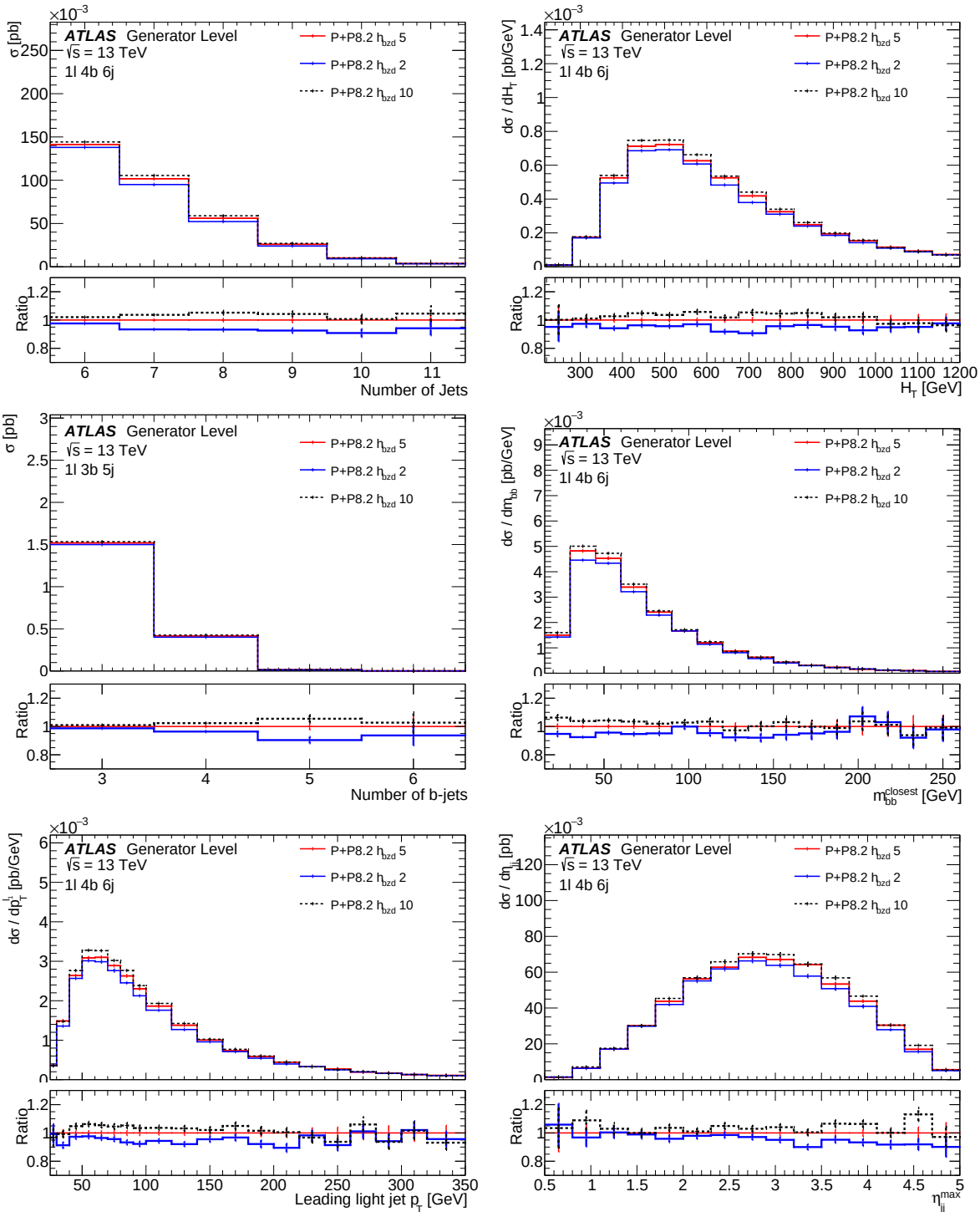

Figure 5: Kinematic distributions with different h_{bzd} settings using predictions with decayed top quarks in the 1l 4b 6j region. The ratio is computed with respect to the P+P8.2 $h_{\mathrm{bzd}} = 5$ prediction. Jet multiplicity (upper left), H_{T} (upper right), b-jet multiplicity (middle left), m_{bb} (middle right), light jet p_{T} (lower left), $\eta_{\mathrm{jj}}^{\mathrm{max}}$ (lower right).

4.3 Decay handling and spin correlation setup

The decay of the top quarks is calculated in POWHEG and in MadSpin [20] preserving spin correlations. Both codes make predictions at the same level of precision and use similar calculations but differ in the details. Both can be used for the production of $t\bar{t}b\bar{b}$ events. For the new set-up, POWHEG would be preferred for technical reasons to maintain compatibility with the generation of the $t\bar{t}H$ process, which also uses POWHEG for the decay handling of the top quarks.

To validate this change, the predictions of the POWHEG and MadSpin calculations are compared. Similar results were found for most of the distributions. Figure 6 shows the azimuthal opening angle between leptons in dileptonic $t\bar{t}b\bar{b}$ events which is known to be particularly sensitive to the spin correlations of the top quarks. A small difference is observed, as was also observed for the inclusive top quark pair production shown in Ref. [29]. MadSpin predicts a slightly higher cross-section at small opening angles, with the largest difference of 10% observed at $\Delta\phi(l^+, l^-)/\pi < 0.3$ in 4 b-jet events. Due to the similarity of predictions and no data to discriminate between them, POWHEG was chosen for the decay.

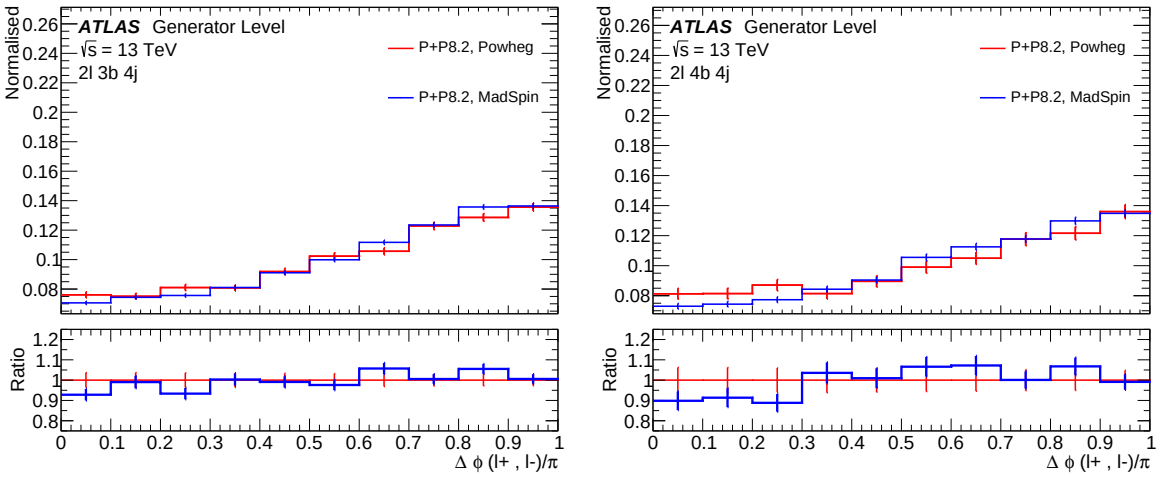

Figure 6: MadSpin and POWHEG predictions of the azimuthal opening angle of the two leptons ($\Delta\phi(l^+, l^-)/\pi$) in the dilepton channel using predictions with decayed top quarks in the dilepton regions as defined in Section 3.1. The predictions are compared in typical $t\bar{t}H$ signal regions. The new default is marked in red. The plot on the left displays events with three b-jets and events with four b-jets are shown on the right.

4.4 Reduction of negative-weighted events

As the name of the generator "Positive Weight Hardest Emission Generator" [26] indicates the POWHEG method aims at predictions that produce mainly positive-weighted events. However, a small fraction of negative-weighted events caused by the subtraction terms in the NLO calculations may still exist. The number of negative-weighted events increases with increasing α_s which occurs e.g. when lowering the renormalisation scale. In order to counter this trend and reduce the number of negative weighted events, a so-called folding method was introduced in the MC integration as described in Ref. [30].

The POWHEG program calculates the integration in the radiation variables y, ξ and ϕ (defined in Ref. [30]) and correspondingly the number of foldings can be steered by the folding parameters foldy, foldcsi and foldphi [31, 32]. Since the increase in the number of foldings will lead to an increase in the production time a compromise between the improvement in the fraction of negative-weight events and the increase in production time has to be found. In Table 7 different fold parameter settings are shown with the resulting fraction of negative-weight events and the increase in production time compared to the setting 551. Furthermore, the fraction of negative-weight events for the scale down variation of $0.5 \cdot \mu_R$, $0.5 \cdot \mu_F$ is shown which is the highest among all scale variations. The 555 setup is found to be optimal since it reduces the fraction of negative-weight events by almost a factor of two with an acceptable increase of production time of 55%.

Table 7: Fraction of negative-weight events and the relative increase in production time for l+jets $t\bar{t}b\bar{b}$ samples as function of the folding parameter settings.

Name	foldcsi	foldy	foldphi	neg. fraction (nominal)	neg. frac. (scale down variation)	prod. time increase
551	5	5	1	9.7%	47.1%	-
552	5	5	2	9.1%	46.2%	10%
555	5	5	5	5.2%	33.1%	55%
1055	10	5	5	4.1%	32.7%	140%

5 Optimising the matching with PYTHIA8

Studies on the settings in the PYTHIA8 code affecting the matched prediction of POWHEG+PYTHIA8 are presented in this section. The section is structured as follows: Section 5.1 discusses the p_{T} definition used for the calculation of the hardness criterion for the shower veto and then studies the effect of p_{T} definition on the particle level predictions. Section 5.2 demonstrates the effect of the shower recoil model on the predictions.

5.1 Transverse momentum definition used for matching

The goal in matching PYTHIA8 to the ME calculated by POWHEG is to cover the full radiation phase space but avoid double counting. To achieve this, the recommended method of vetoed shower is used. PYTHIA8 generates radiation over the full phase space but emissions covered by POWHEG are vetoed. This is possible since both POWHEG and PYTHIA8 are based on a combined evolution of initial-state shower (ISR) and final-state shower (FSR) in p_{T}-related "hardness" variables.
Technically the matching is implemented such that POWHEG provides Born-type events (with no emissions) and Real-type events (with an additional parton), as well as the hardness criterion with its value used to separate Born-type events from Real-type events. PYTHIA8 generates emissions using its PYTHIA-hardness ordering and the POWHEG-hardness criterion is used to veto emissions above the POWHEG hardness value of the input event.
Complications arise, however since the hardness definitions in POWHEG and PYTHIA8 differ. To calculate the POWHEG hardness criterion different p_{T} definitions are available in the PYTHIA8 code, separately for ISR and FSR emissions, and can be steered by the parameter POWHEG:pTdef referred to as $pTdef$ in the following. The first option uses the POWHEG ISR p_{T} definition for both ISR and FSR ($pTdef = 0$), i.e. the p_{T} of the emitted parton is calculated with respect to the axis of the radiating parton [3]. The second option uses also the POWHEG ISR p_{T} for the ISR emissions but the d_{ij} p_{T} definition for FSR ($pTdef = 1$) where d_{ij} p_{T} is calculated as the relative p_{T} of one of the particles with respect to the axis of the second after the emission. Finally using $pTdef = 2$ it is possible to set in PS the PYTHIA8 ISR and FSR p_{T} definitions for the respective cases. Following the logic of the vetoed shower, the authors recommend to use the POWHEG definitions as steered by $pTdef = 1$.

The effect of the different p_{T} definitions in the shower veto was evaluated using the analysis with decayed top quarks described in Section 3. No significant difference between the three definitions was observed in the kinematic distributions, as demonstrated by a few examples in Figure 7, however the jet multiplicity is slightly reduced if the PYTHIA8 p_{T} definition is used ($pTdef = 2$). Based on these results, it was decided to use $pTdef = 1$ as default value.

[3] corresponding to the PYTHIA8 default setting $POWHEG : pTemt = 0$

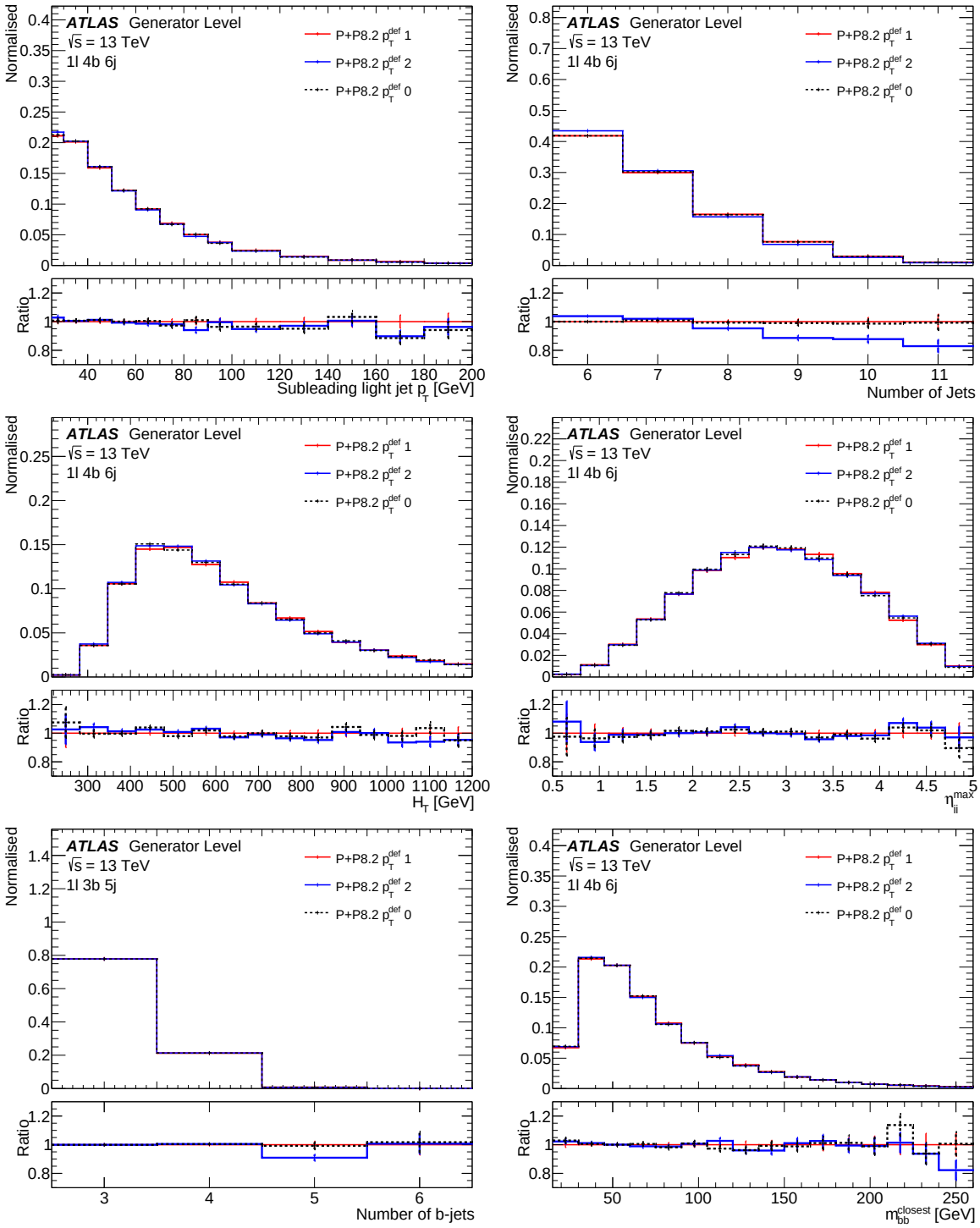

Figure 7: Kinematic distributions with different $pTdef$ settings using predictions with decayed top quarks in the 1l 4b 6j region as defined in Section 3.1. The chosen default is marked in red and taken as the reference for the ratio. Subleading light jet p_{T} (upper left), number of jets (upper right) , H_{T} (middle left), $\eta_{\mathrm{jj}}^{\mathrm{max}}$ (middle right), number of b-jets (lower left) and m_{bb} (lower right) .

5.2 Treatment of initial state shower recoil

In order to conserve momentum for QCD radiation in the parton shower algorithm, different recoil settings are implemented in PYTHIA8 [33]. The timelike FSR is based on dipole-style recoils, where one single parton takes the full recoil of a branching. Two options exist for the spacelike ISR shower: the first option uses a global recoil where the recoil of an ISR emission is taken by the whole final state[4]. This is the default in PYTHIA8. The second option uses a dipole recoil in which only one final-state parton takes the recoil of an emission. The global recoil is well motivated for cases where the underlying hard process does not involve color flow between the initial and the final state, such as deep inelastic scattering (DIS) or where the color flow is only between the initial state quark and the final state quark like in VBF Higgs production [34]. For the process of $t\bar{t}b\bar{b}$ no clear theoretical preference for either case can be formulated [35]. Therefore, both options for spacelike ISR showering will be considered.

In addition to the usual observables discussed above, a special variable is constructed with sensitivity to the recoil. If the recoil of the system is dominantly absorbed by one of the final state partons, one would expect an angular difference of $\Delta\phi \approx \pm\pi$ between this parton and the sum p_{T} of the final state partons. This is implemented as follows:

$$\vec{p}_{\mathrm{rec}} = \sum_{\mathrm{i}=t,\bar{t},b,\bar{b}} \vec{p}_{\mathrm{T_i}}$$
$$\Delta\phi_{\mathrm{rec}} = \Delta\phi(\vec{p}_{\mathrm{rec}}, \vec{p}_{\mathrm{X}})$$

where the sum runs over top quark, antitop quark, b-quarks and leading additional jets if existent. X refers to the partons in the event.

Figure 8 shows the $\Delta\phi_{rec}$ distribution where $\vec{p}_{\mathrm{X}}$ is taken to be the momentum of the leading top or the leading additional b-jet , for the two options using analysis with stable top quarks in the $t\bar{t}b\bar{b}$ region. The recoil is absorbed by the leading top quark as a peak structure at the values of $\pm\pi$ is visible while the $\Delta\phi_{rec}$ to the leading b-jet is rather flat. It appears that the distributions are not sensitive to the particular recoil scheme used in PYTHIA8 as no significant difference is observed between the different schemes neither for the leading top quark nor for the additional b-jet . However, the additional light jets are observed to be softer for the dipole recoil, resulting in a slightly softer H_{T} spectrum and significantly fewer events with high light jet multiplicity as shown in Figure 8.

The effects observed with the stable top quark analysis remains visible for the particle level final state. Figure 9 shows the corresponding distributions for the single lepton channel. Kinematic distributions of the leading b-jet are not affected, while light jet p_{T} and H_{T}, tend to be softer for the dipole shower, leading also to significantly fewer events with high jet multiplicity.

In addition to the two POWHEG+PYTHIA8 recoil predictions an alternative prediction using HERWIG7 with the angular ordered PS is shown. The POWHEG+PYTHIA8 with the dipole scheme and POWHEG+HERWIG7 predictions are similar, especially in the jet multiplicity distribution where both have significantly less events with high jet multiplicities in comparison to the POWHEG+PYTHIA8 global recoil prediction. This indicates that the difference between PYTHIA8 and HERWIG7 which is often evaluated as part of the modelling uncertainty might be caused by the shower recoil scheme to a large extend.

[4] The options are steered via SpaceShower:dipoleRecoil "on/off"in the PS

The different settings were also compared to the ATLAS measurements of $t\bar{t}b\bar{b}$ but due to the large uncertainties, the data could not discriminate between them as demonstrated in Figure 10. Therefore, both predictions are considered to estimate the uncertainty on the $t\bar{t}b\bar{b}$ modelling. As it stated in Tab. 6 the previous nominal (global recoil) is kept and a prediction using the dipole recoil is used to estimate the uncertainty of the PS spacelike recoil scheme.

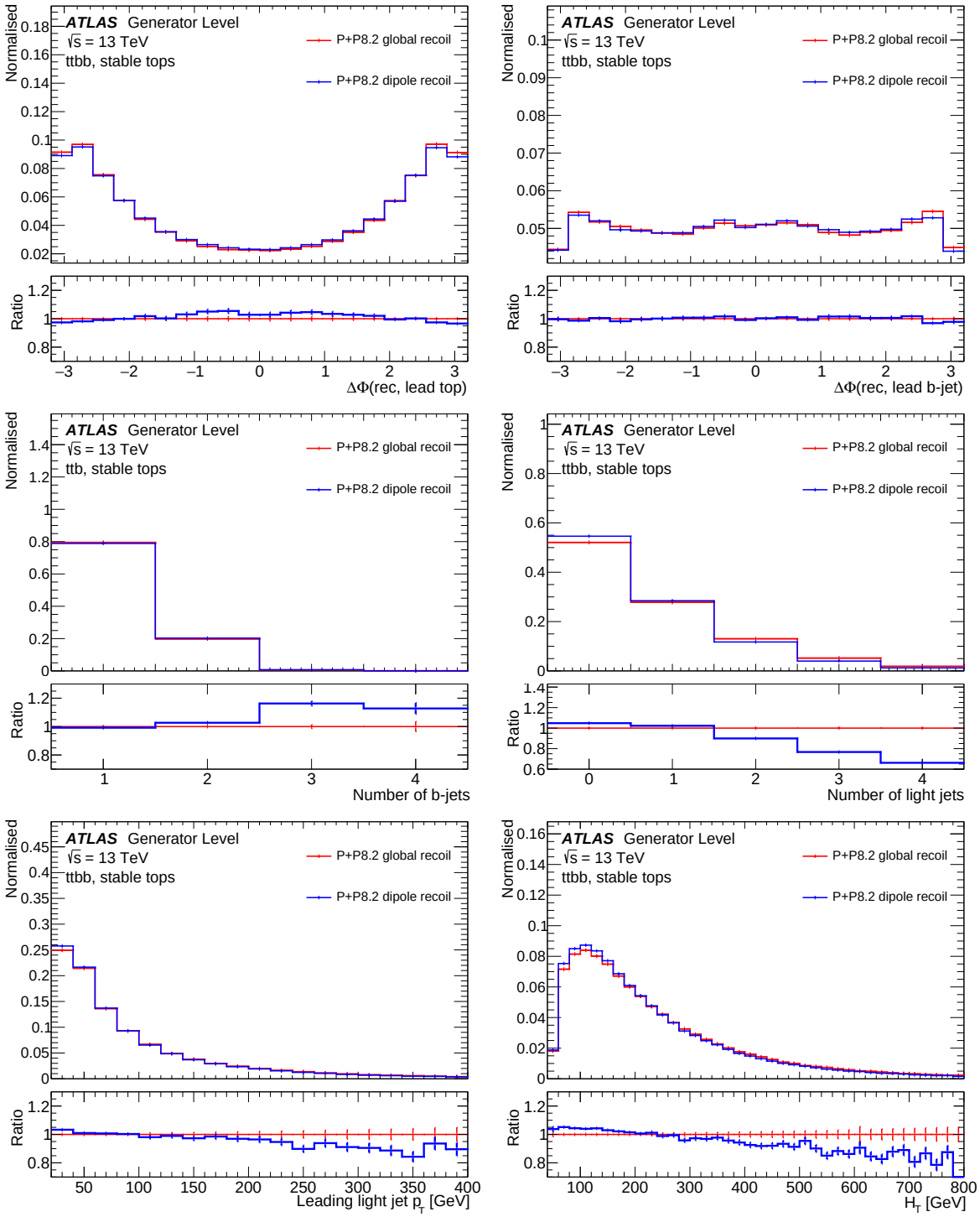

Figure 8: POWHEG+PYTHIA8 predictions with different models for the ISR recoil in using the $t\bar{t}b\bar{b}$ region of the stable top analysis as described in Section 3.2. The ratio is computed with respect to the P+P8.2 global recoil prediction. $\Delta\phi_{rec}$ against the leading top quark (upper left), $\Delta\phi_{rec}$ against the leading additional b-quark (upper right), b-jet multiplicity (middle left), jet multiplicity (middle right), leading light jet p_{T} (lower left), H_{T} (lower right).

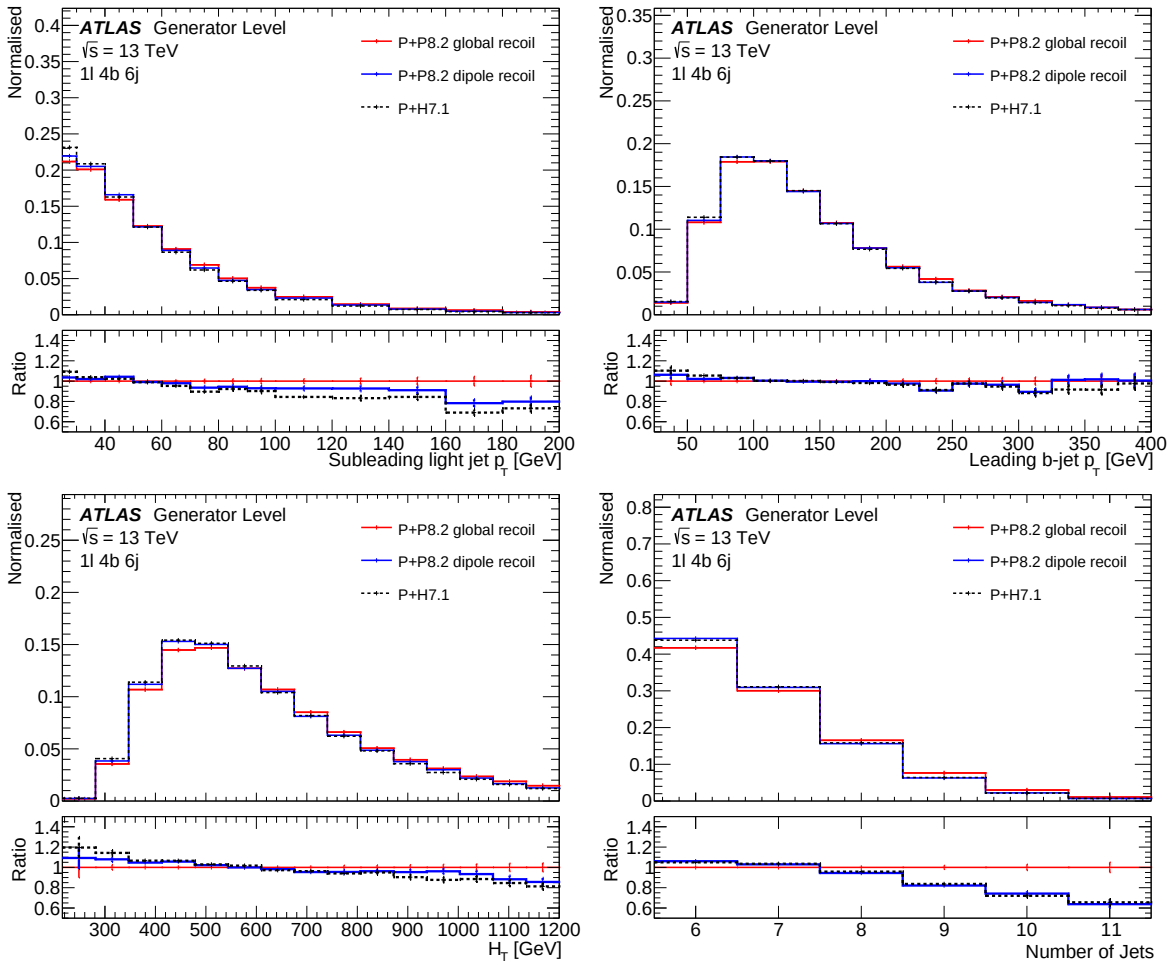

Figure 9: Kinematic distributions comparing predictions of POWHEG+PYTHIA8 with the different recoil schemes are compared to POWHEG+HERWIG7 using predictions with decayed top quarks in the 1l 4b 6j region as defined in Section 3.1. The ratio is computed with respect to the P+P8.2 global recoil prediction. Subleading light jet p_{T} (upper left), leading b-jet p_{T} (upper right), H_{T} (lower left) , jet multiplicity (lower right).

6 Comparisons of generator predictions

The presented studies are leading to an optimised set-up for POWHEG and the matched POWHEG+PYTHIA8 predictions. In the following, scale variations are applied and comparisons to further predictions are made in order to get an estimate of the overall modelling uncertainties of the $t\bar{t}b\bar{b}$ process calculated with at NLO using the 4 FS. The uncertainty due to matching and the choice of MC generator is estimated comparing to SHERPA. The uncertainty related to the choice of PS calculation is estimated by comparing to POWHEG+HERWIG7 and POWHEG+PYTHIA8 with the dipole recoil shower.

Figure 10 shows the predictions compared to $t\bar{t}b\bar{b}$ differential measurements from ATLAS. Good agreement of all predictions with the data is observed. Due to the small differences between the predictions and the large experimental uncertainties, no further constraints beyond the scale setting discussed in Section 4.1 can be derived.

Figures 11-14 show particle level distributions in the kinematic regions typically used for the $t\bar{t}b\bar{b}$ measurement and the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis in the lepton+jets and the dilepton decay channels. The distributions are normalised to unity. Overall a good agreement between generators is observed. The scale uncertainties and differences between predictions are small on observables reconstructed dominantly from b-jets initiated by b-quarks from the ME calculation such as b-jet p_{T}, and m_{bb} . However differences are observed related to additional radiation i.e. light jet p_{T}, H_{T}, jet multiplicity and b-jet multiplicity for more than four b-jets. SHERPA and POWHEG+HERWIG7 differ most in these distributions but nicely enclose POWHEG+PYTHIA8. As expected, the shower variations lead to rather small differences in POWHEG matched predictions as long as the multiplicity is comparable with the number of partons from the ME calculation. At high jet and b-jet multiplicities where partons from PS are contributing, the PS settings significantly influence the predictions. Overall, shape differences of up to ± 10% for observables related to b-jets and up to ± 20% for observables related to additional jet production are observed between the different $t\bar{t}b\bar{b}$ predictions in the bulk part of the phase space with increased differences in the tails of some distributions which however contribute very little to the observed number of events.

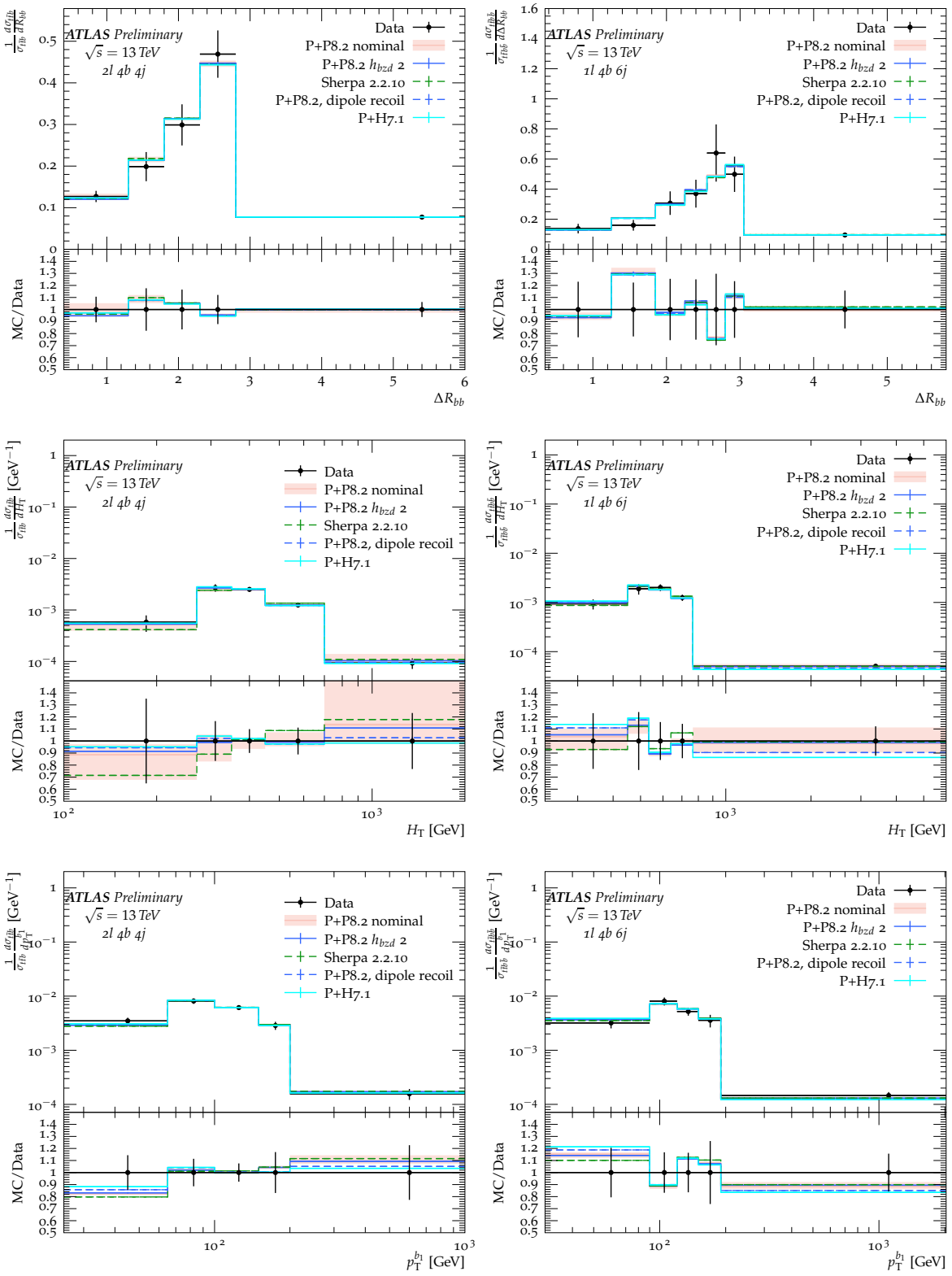

Figure 10: Comparison of the POWHEG predictions with different PS settings and SHERPA to the ATLAS $t\bar{t}b\bar{b}$ measurement [16] . On the left (right) side comparisons in the di-lepton ($l + jets$) top quark decay channel are shown. ΔR_{bb} (upper row), H_{T} (middle row) and leading b-jet p_{T} (lower row). Scale variations are included in form of a band spanned by simultaneously varying μ_{R} μ_{F} up and down by a factor of 2 compared to the nominal $t\bar{t}b\bar{b}$ prediction

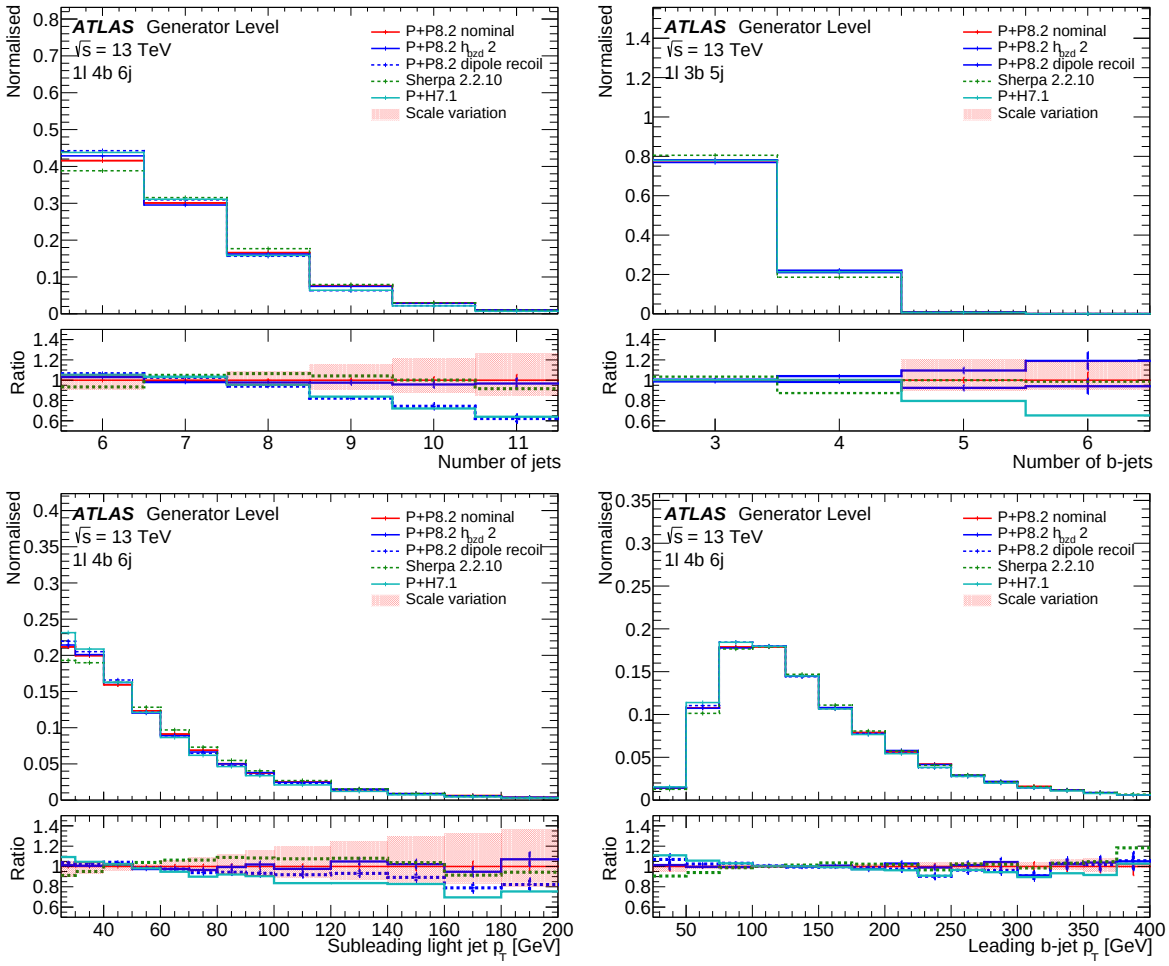

Figure 11: Comparisons of $t\bar{t}b\bar{b}$ predictions in the l+jets channel using the analysis with decayed top quarks in the 1l 4b 6j region as described in Section 3.1. The ratio is computed with respect to the P+P8.2 global recoil prediction. All distributions are normalised to unit area. Number of jets (upper left), number of b-jets (upper right), subleading light jet p_{T} (lower left), leading b-jet p_{T} (lower right). Scale variations are included in form of a band spanned by simultaneously varying μ_{R} μ_{F} up and down by a factor of 2 compared to the nominal $t\bar{t}b\bar{b}$ prediction.

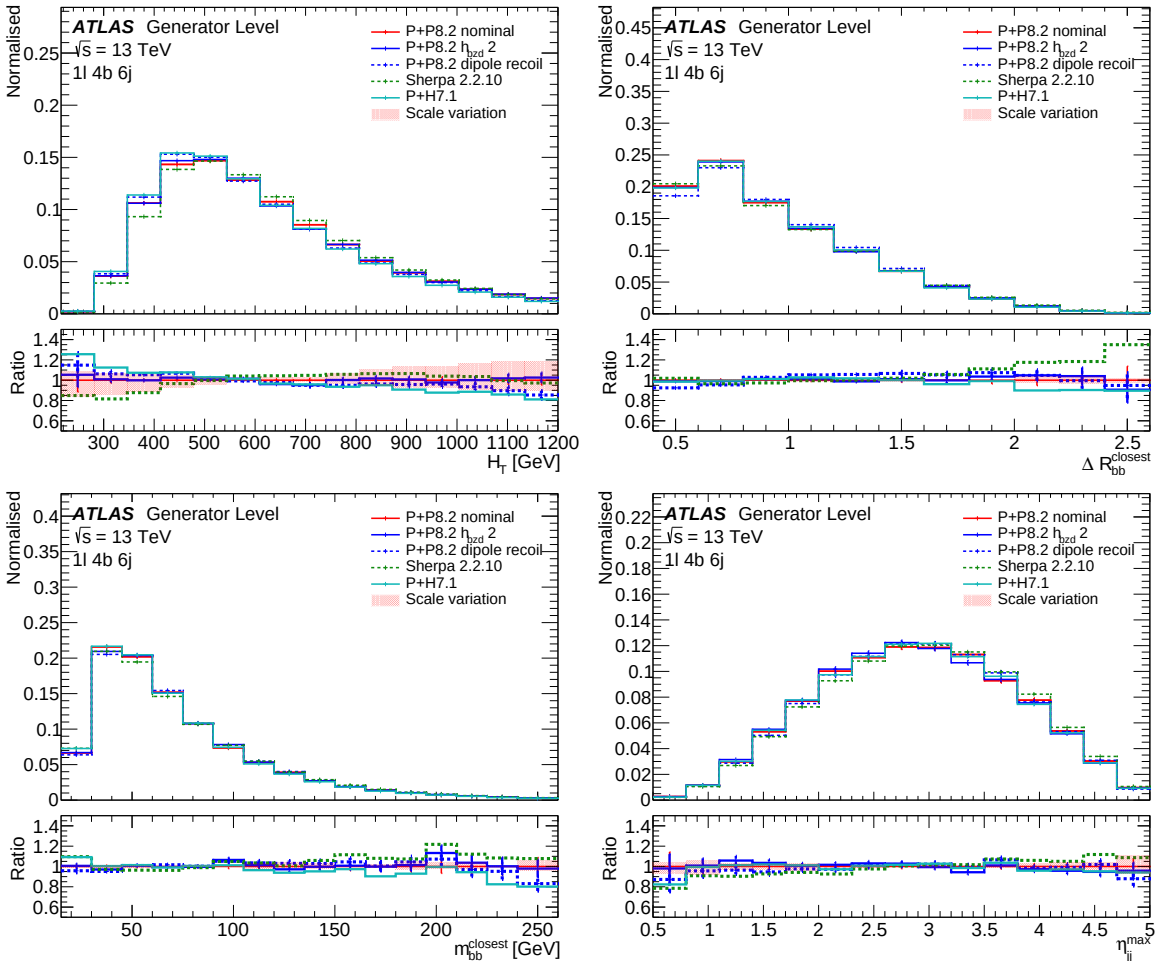

Figure 12: Comparisons of $t\bar{t}b\bar{b}$ predictions in the l+jets channel using the analysis with decayed top quarks in the 1l 4b 6j region as described in Section 3.1. The ratio is computed with respect to the P+P8.2 global recoil prediction. All distributions are normalised to unit area. H_{T} (upper left), ΔR_{bb} (upper right), m_{bb} (lower left), $\eta_{\mathrm{jj}}^{\mathrm{max}}$ (lower right). Scale variations are included in form of a band spanned by simultaneously varying μ_{R} μ_{F} up and down by a factor of 2 compared to the nominal $t\bar{t}b\bar{b}$ prediction

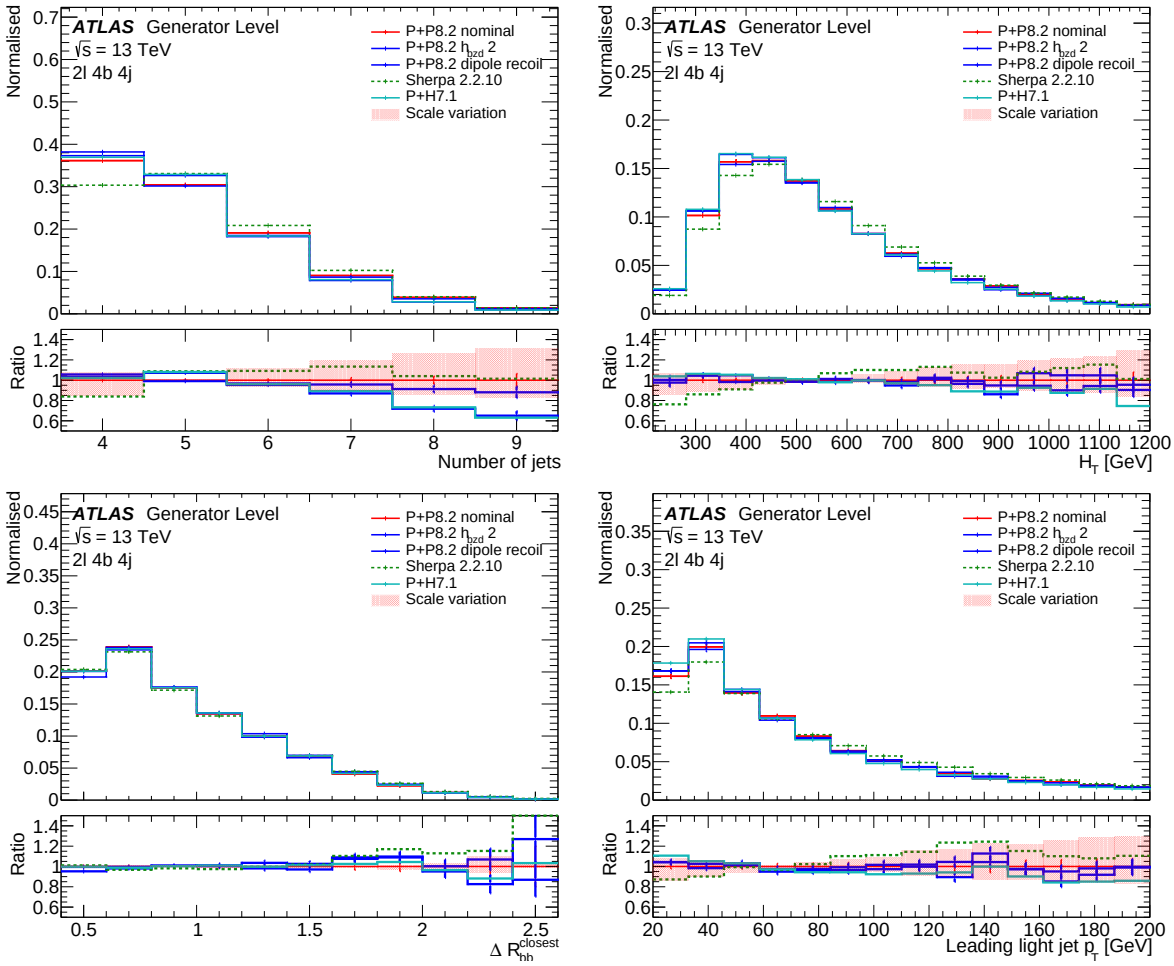

Figure 13: Comparisons of $t\bar{t}b\bar{b}$ predictions in the dilepton channel using the analysis with decayed top quarks in the 2l 4b 4j region as described in Section 3.1. The ratio is computed with respect to the P+P8.2 global recoil prediction. All distributions are normalised to unit area. Number of jets (upper left), H_{T} (upper right), ΔR_{bb} (lower left), leading light p_{T} (lower right). Scale variations are included in form of a band spanned by simultaneously varying μ_{R} μ_{F} up and down by a factor of 2 compared to the nominal $t\bar{t}b\bar{b}$ prediction

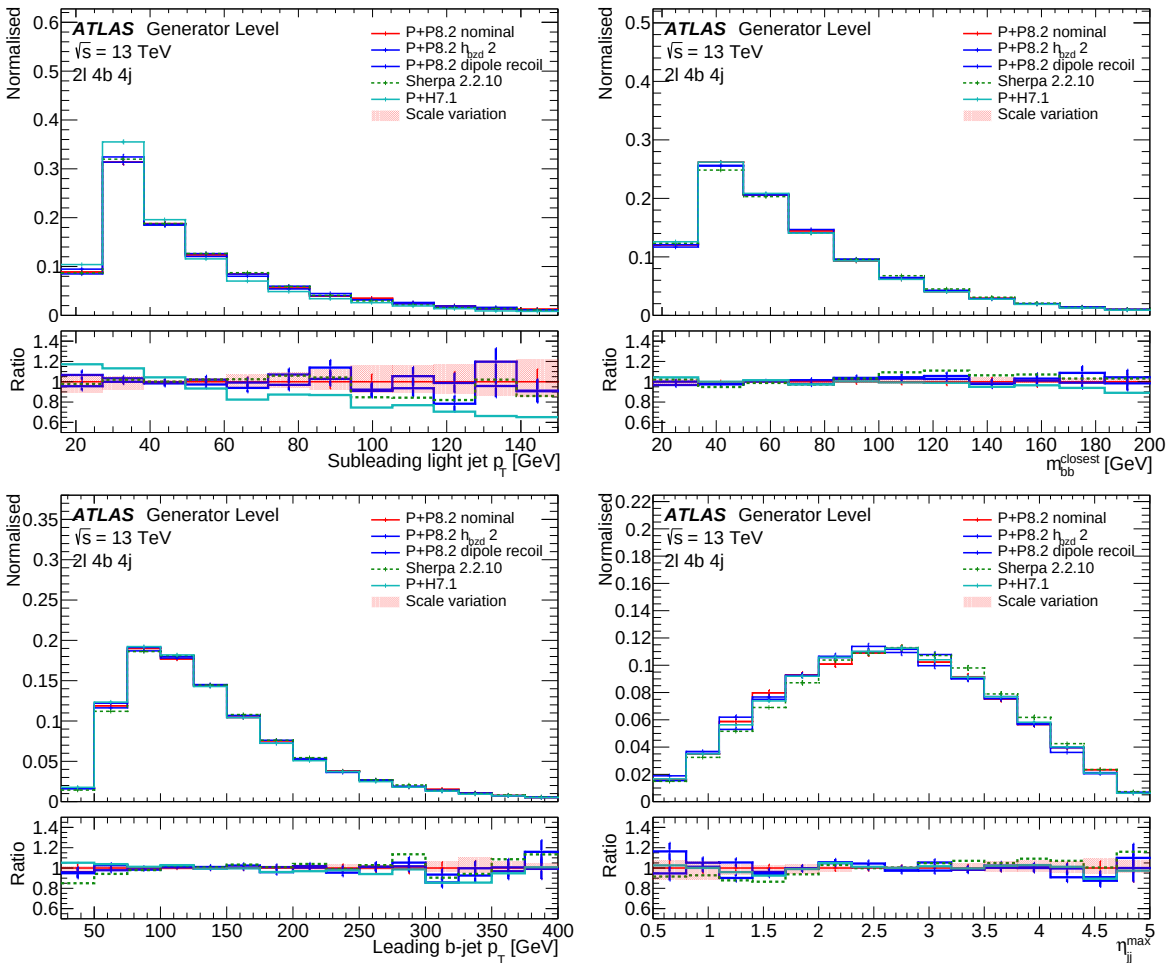

Figure 14: Comparisons of $t\bar{t}b\bar{b}$ predictions in the dilepton channel using the analysis with decayed top quarks in the 2l 4b 4j region as described in Section 3.1. The ratio is computed with respect to the P+P8.2 global recoil prediction. All distributions are normalised to unit area. Subleading light jet p_{T} (upper left), m_{bb} (upper right), leading b-jet p_{T} (lower left), $\eta_{\mathrm{jj}}^{\mathrm{max}}$ (lower right). Scale variations are included in form of a band spanned by simultaneously varying μ_{R} μ_{F} up and down by a factor of 2 compared to the nominal $t\bar{t}b\bar{b}$ prediction

7 Conclusion and summary

This note presents a new POWHEG+PYTHIA8 setup to simulate $t\bar{t}b\bar{b}$ events at the LHC with a centre-of-mass energy of 13 TeV. The spin correlation and decay handling has been changed, using now the default method in POWHEG. Parameter settings in POWHEG and in its matching to PYTHIA8 were optimised based on theoretical arguments and on comparisons to data resulting in the following proposed setup: a reduction of the renormalisation scale to improve the agreement with data and to reduce k-factors, a set of folding parameters that reduce the fraction of negative-weight events by a factor of two, a change of the POWHEG matching parameter ($pTdef$) and a change in the POWHEG damping parameter h_{bzd}.

Based on the findings in the presented studies we also recommend to add in the estimate of modelling uncertainties predictions of POWHEG matched to PYTHIA8 with the dipole recoil shower and variations of the POWHEG damping parameter (h_{bzd}) in the estimate of modelling uncertainties. Finally, a comprehensive study of all modelling uncertainties of the $t\bar{t}b\bar{b}$ process is shown comparing the new POWHEG+PYTHIA8 prediction and its scale variations, the POWHEG damping parameter and the PYTHIA8 and HERWIG parton showers to SHERPA which revealed differences larger than the scale variations in some observables.

References

[1] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD corrections to pp —> t anti-t b anti-b + X at the LHC*, (2009), URL: https://arxiv.org/abs/0905.0110 (cit. on p. 2).

[2] M. V. Garzelli, A. Kardos, and Z. Trócsányi, *Hadroproduction of $t\bar{t}b\bar{b}$ final states at LHC: predictions at NLO accuracy matched with Parton Shower*, (2015), URL: https://arxiv.org/abs/1408.0266 (cit. on p. 2).

[3] 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. C **78** (2018) 502, arXiv: 1802.00426 [hep-ph] (cit. on pp. 2, 3, 6–10).

[4] ATLAS Collaboration, *Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, CERN-EP-2021-202, 2021, URL: https://arxiv.org/abs/2111.06712 (cit. on pp. 2, 3, 6, 7).

[5] ATLAS Collaboration, *Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, CERN-EP-2021-075, 2021, URL: https://cds.cern.ch/record/2773682 (cit. on pp. 2, 6).

[6] ATLAS Collaboration, *Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2020-023, 2020, URL: https://cds.cern.ch/record/2730443 (cit. on p. 2).

[7] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 2).

[8] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 2).

[9] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 2).

[10] 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. B **734** (2014) 210, arXiv: 1309.5912 [hep-ph] (cit. on pp. 2, 3).

[11] 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. 2, 3).

[12] A. Buckley et al., *Rivet user manual*, Comput. Phys. Commun. **184** (2013) 2803, arXiv: 1003.0694 [hep-ph] (cit. on p. 2).

[13] 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. 2).

[14] A. Ryd et al., *EvtGen: A Monte Carlo Generator for B-Physics*, (2005), URL: https://evtgen.hepforge.org/doc/EvtGenGuide.pdf (cit. on p. 3).

[15] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, (2012), URL: https://arxiv.org/abs/1111.1220 (cit. on p. 3).

[16] ATLAS Collaboration,
Measurements of fiducial and differential cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector,
ATLAS-CONF-2018-029, 2018, URL: https://cds.cern.ch/record/2628772
(cit. on pp. 3, 5, 8, 24).

[17] 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. 3).

[18] 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. 3).

[19] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*,
Journal of High Energy Physics **2004** (2004) 040–040, ISSN: 1029-8479,
arXiv: arXiv:hep-ph/0409146 [hep-ph] (cit. on p. 6).

[20] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk,
Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations,
JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph] (cit. on pp. 6, 15).

[21] ATLAS Collaboration, *Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, (),
URL: https://arxiv.org/abs/2007.14858 (cit. on pp. 6, 7).

[22] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD Corrections to Top Anti-Top Bottom Anti-Bottom Production at the LHC: 2. full hadronic results*, JHEP **03** (2010) 021,
arXiv: 1001.4006 [hep-ph] (cit. on p. 6).

[23] 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. B **734** (2014) 210,
arXiv: 1309.5912 [hep-ph] (cit. on p. 6).

[24] F. Buccioni, S. Kallweit, S. Pozzorini, and M. F. Zoller,
NLO QCD predictions for $t\bar{t}b\bar{b}$ production in association with a light jet at the LHC,
JHEP **12** (2019) 015, arXiv: 1907.13624 [hep-ph] (cit. on p. 7).

[25] S. Pozzorini, "Study of different NLO MC frameworks: summary of results",
Experimental and Theoretical studies of tt+bb and ttW, 2020,
URL: https://indico.cern.ch/event/964993/contributions/4075701/attachments/2129120/3585126/ttbb-status-23oct20.pdf (cit. on p. 7).

[26] 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 pp. 9, 16).

[27] 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. 9).

[28] S. Pozzorini, private communication (cit. on p. 10).

[29] ATLAS Collaboration, *Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s}$ = 13 TeV using pp collisions in the ATLAS detector*, Eur. Phys. J. C **80** (2020) 754,
arXiv: 1903.07570 [hep-ex] (cit. on p. 15).

[30] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi,
W^+W^+ plus dijet production in the POWHEGBOX, Eur. Phys. J. C **71** (2011) 1670,
arXiv: 1102.4846 [hep-ph] (cit. on p. 16).

[31] 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. 16).

[32] P. Nason, *MINT: A Computer program for adaptive Monte Carlo integration and generation of unweighted distributions*, (2007), arXiv: 0709.2085 [hep-ph] (cit. on p. 16).

[33] B. Jager, A. Karlberg, and J. Scheller, *Parton-shower effects in electroweak $WZjj$ production at the next-to-leading order of QCD*, Eur. Phys. J. C **79** (2019) 226, arXiv: 1812.05118 [hep-ph] (cit. on p. 19).

[34] B. Jäger, A. Karlberg, S. Plätzer, J. Scheller, and M. Zaro, *Parton-shower effects in Higgs production via Vector-Boson Fusion*, Eur. Phys. J. C **80** (2020) 756, arXiv: 2003.12435 [hep-ph] (cit. on p. 19).

[35] T. Sjöstrand, private communication (cit. on p. 19).