Hadroproduction of $t\bar{t}b\bar{b}$ final states at LHC: predictions at NLO accuracy matched with parton shower

M.V. Garzelli,[a] **A. Kardos**[a] **and Z. Trócsányi**[b]

[a]*MTA-DE Particle Physics Research Group, University of Debrecen,*
H-4010 Debrecen P.O. Box 105, Hungary

[b]*Institute of Physics and MTA-DE Particle Physics Research Group, University of Debrecen,*
H-4010 Debrecen P.O. Box 105, Hungary

E-mail: garzelli@mi.infn.it, kardos.adam@science.unideb.hu,
Zoltan.Trocsanyi@cern.ch

ABSTRACT: We present predictions for the hadroproduction of $t\bar{t}b\bar{b}$ final states at the LHC with collision energies $\sqrt{s} = 8\,\text{TeV}$ and $14\,\text{TeV}$ at NLO accuracy matched with parton shower, as obtained with PowHel + PYTHIA. We quantify the effects of parton shower and hadronization. We find these are in general moderate except the effect of the decay of heavy particles, which can modify significantly some distributions, like that of the invariant mass of the two leading b-jets. We also show kinematic distributions obtained with cuts inspired to those recently employed by the CMS collaboration. For these predictions, we present the theoretical uncertainty bands, related to both scale and PDF variations. We find that these uncertainties are only moderately affected by the change of the collision energy.

Contents

1 Introduction

The $t\bar{t}b\bar{b}$ hadroproduction process was one of the first four-parton hadroproduction processes studied a few years ago in order to test and show the capabilities of new unitarity-inspired methods for evaluating one-loop amplitudes, as well as the potential of new developments in the more traditional tensor-reduction approach. Pioneering works in this direction were performed by Bredenstein et al. [1, 2], as well as by the `HELAC-NLO` collaboration [3]. In both cases the role of next-to-leading-order (NLO) QCD corrections was shown, by considering rather loose systems of cuts. The agreement between predictions obtained by different methods was taken as a prove of the correctness of the new conceptual developments, that were subsequently applied successfully to other multiparticle hadroproduction processes, reaching a level of unforeseen complexity [4–7].

A further application was represented by the study of $t\bar{t}b\bar{b}$ production at the TeVatron [8]. Nowadays, a renewed interest for this process has arisen, driven not only by further theoretical developments, including the possibility of matching complex NLO QCD calculations to parton shower (PS) approaches, but also by the requests of the experimental collaborations working at LHC, that need input to reduce the uncertainties on the estimate of the $t\bar{t}b\bar{b}$ contribution to the total background in their searches for the Higgs boson in the $t\bar{t}H$ channel, with the H boson decaying in a $b\bar{b}$ pair [9–12]. The theoretical uncertainty quoted in recent experimental analyses on the contribution of this irreducible non-resonant background component amounts to a few tens percent ($\sim 50\%$ in refs. [11, 12]). We expect that theoretical estimates with the best possible theoretical accuracy within present reach of the most advanced event generators can help reduce this uncertainty. Thus, one of the aims of this paper is to present predictions for differential cross sections at NLO QCD accuracy matched to PS and hadronization, resulting in predictions at the hadron level, using cuts similar to those employed by the experimental collaborations.

First predictions concerning $t\bar{t}b\bar{b}$ computations with NLO QCD accuracy matched with PS were presented by our group in refs. [13, 14], making use of `PowHel`. More recently,

$t\bar{t}b\bar{b}$ predictions obtained by using the MC@NLO matching algorithm [15], as encoded in SHERPA [16], and massive b-quarks, with a fixed pole mass $m_b = 4.75\,\mathrm{GeV}$, have been reported in ref. [17], using 1-loop amplitudes computed by OpenLoops [18]. Our computation differs from that presented in ref. [17] in the following aspects: we use (i) massless b-quarks in the generation of NLO matrix elements, (ii) a different matching algorithm (POWHEG [19, 20]), and (iii) a different Shower Monte Carlo program (PYTHIA).

As pointed out in ref. [17], the finite mass of the b-quarks allows for computing the contribution of $t\bar{t}jj$ events with the jets originating from two gluons that both separately split into a collinear $b\bar{b}$ pair. Each pair itself becomes part of one b-jet (as the b-quarks are sufficiently collinear to be unresolved by the jet algorithm), so the final state contains a $t\bar{t}$-quark pair in association with two b-jets. Strictly speaking, such contributions are higher order in perturbation theory and are not included in our computation with NLO+PS accuracy, where collinear $b\bar{b}$ pairs can only be produced by gluon splitting in the PS. On the other hand, in ref. [17], one of those $g \to b\bar{b}$ splittings is included in the real radiation matrix elements (at LO accuracy for these contributions), while the other is generated by the parton shower, and such double $g \to b\bar{b}$ splittings were found to give a significant contribution to the $t\bar{t}+2$ b-jet final states when the t-quarks are kept stable.

This paper represents an extension of our previous studies, relying on a more advanced PowHel implementation, on the basis of recent POWHEG-BOX developments, and includes a detailed study of the theoretical uncertainties affecting our results in the most general case, including top decay and the full SMC chain up to hadronization and hadron decay.

2 Method

Predictions presented in this paper were obtained using events generated by PowHel [21], and stored in Les Houches event (LHE) files [22]. Our computational framework, PowHel, is an event generator that uses the HELAC-NLO codes [23] (HELAC-1LOOP [24] and HELAC-Dipole [25]) for the computation of the amplitudes required as input by the POWHEG-BOX [26]. We use the latter for matching NLO QCD computations with PS, according to the POWHEG method, generating events stored in the Les Houches format (LHE). Details of event generation for $t\bar{t}b\bar{b}$ final states were presented in ref. [13], where thorough checks of the correctness of the event files were also discussed. In particular, we treated the b-quarks as massless, and used five massless flavours, but neglected the small contribution to the cross section from the b-quarks in the initial state. Feynman graphs containing $g \to b\bar{b}$ splittings are included in our computations. Such splittings are singular when the massless b-quarks become collinear. To avoid this divergence, we employ a generation cut on the invariant mass of the $b\bar{b}$ pair (has to be above $2\,\mathrm{GeV}$), which is chosen such that it is harmless when physical cuts are also employed [13]. In practice, this means that the physical cuts must be chosen such that the invariant mass of b-jets should be significantly larger than $2\,\mathrm{GeV}$. We tested that distributions are independent of this generation cut if m_{bb} is above about $20\,\mathrm{GeV}$. We set the mass of the t-quark to $m_t = 173.2\,\mathrm{GeV}$ for the events at $8\,\mathrm{TeV}$ and to $m_t = 172.6\,\mathrm{GeV}$ for those at $14\,\mathrm{TeV}$. The latter choice was made in the first calculation at $14\,\mathrm{TeV}$ [3], against which we checked our predictions at NLO accuracy in ref. [13].

Subsequent radiation emissions on top of the LHEs are generated by shower Monte Carlo (SMC) programs. SMCs can be further used to generate hadronization and hadron decay, as well as the decay of the heavy elementary particles in the narrow width approximation (NWA). For making predictions in this paper, we used the `PYTHIA` SMC code [27], in the Fortran version 6.4.28. However, other SMCs can be used as well to further shower the events generated by `PowHel`.

Following ref. [13], we fix the default value of the renormalization and factorization scales equal to $\mu_R = \mu_F = \mu_0 = H_T/2$. Here H_T is the sum of the transverse masses of partons in the final state, a fairly standard choice for final states with high multiplicity [28], using the underlying Born kinematics. Then we consider their simultaneous variation by a factor of two around this value, leading to scale uncertainty bands corresponding to scales in the range $[\mu_0/2, 2\mu_0]$. We also consider μ_R and μ_F variations separately in this range and we verify that the uncertainties associated to off diagonal variations $(\mu_R, \mu_F/2)$, $(\mu_R/2, \mu_F)$, $(\mu_R, 2\mu_F)$, $(2\mu_R, \mu_F)$ lie within the bands of diagonal ones. Our scale uncertainty bands represent the envelope coming from these variations, leaving out the antipodal choices $(\mu_R/2, 2\mu_F)$ and $(2\mu_F, \mu_R/2)$ according to the prescription of ref. [29].

We used the `CT10NLO`, `MSTW2008NLO` and `NNPDF` sets of PDFs with 2-loop running α_S and 5 active flavours, as available in the `LHAPDF` interface [30], drawing a PDF uncertainty band as the envelope of the results corresponding to the three central values of these PDF sets. For practical reasons, due to the intensive CPU requirements of our calculations, we neglected uncertainties coming from considering the statistical variations around each central PDF distribution. We expect that differences between different PDF sets are more pronounced than differences within each PDF set, as already found for the hadroproduction of $t\bar{t}H$ final states in the HXSWG reports [14, 31].

As discussed in ref. [13], suppression factors are used in order to suppress the generation of the events in those regions of the phase space that are expected to be less relevant from the experimental point of view, considering the typical cuts applied in the experimental analyses. These factors however, lead to events characterized by a wide weight distribution, extending without any explicit limit both in the negative and in the positive weight region. This may lead to spikes in the kinematic distributions. We expect these spikes would disappear with significantly more statistics, but the present capabilities of computer resources limit the total number of events we can generate to several millions.[1] We thus implemented an automated spike-elimination procedure applied after SMC, relying on the fact that the same LHE can lead to different SMC emissions, i.e. can populate different bins of the final differential distributions at the hadron level, depending on the random number sequence in the SMC generator.

3 Phenomenology

Using `PowHel` one can make predictions at four different stages in the evolution of the final state: (i) at the parton level using NLO accuracy, (ii) from the pre-showered POWHEG

---

[1]Preparation for event generation requires about 6000 CPU hours, while one event needs on average about 90 seconds on a single recent CPU.

simulation (referred to Les Houches events, or LHEs), formally at the NLO accuracy, (iii) after decay of the heavy particles, (iv) at the hadron level after full SMC. A possible alternative to the last stage is (iv') after PS switching off hadronization and hadron decay, but in this case the heavy quarks are automatically kept stable by PYTHIA. Utilizing these options, we study the effect of these various stages of event evolution on several differential distributions, before making predictions at the hadron level.

3.1 SMC effects

As the number and type of particles are very different after the different stages of event evolution, selection cuts employed at the different stages may change the cross section significantly. Thus, to understand the effect of the PS and hadronization on the shapes of the distributions without changing normalization, we first performed a phenomenological analysis where the cuts are applied on the LHEs, whereas distributions are plotted after different stages of the evolution. This way we can ensure that the selection cuts do not introduce a bias, that is as if we used an inclusive event sample. The cuts applied on jets formed from the LHEs, reconstructed using the $k_\perp$ algorithm [32, 33] with $R = 0.4$, as implemented in `FastJet 3.0.6` [34] are the following: we require at least one b- and one $\bar{b}$-jet with (a) $p_T > 20\,\text{GeV}$, (b) $|\eta| < 2.5$, and (c) $m_{b\bar{b}} > 100\,\text{GeV}$, as relevant for Higgs studies. We do not apply cuts on b-jets emerging from the decay of top quarks, implemented at LO accuracy in the NWA in the SMC, which neglects spin correlations.

In figures 1–6 we present the comparisons of distributions at different stages of event evolution. Each plot contains five predictions, corresponding to the list (i–iv') presented in the introduction of this section.

First we study the effect of the PS by comparing differential distributions from the LHEs and after PS, keeping heavy particles stable. The corresponding predictions are marked as 'LHE' and 'PS'. In the middle section of each plot we also show the ratio PS/LHE of these predictions together with the ratio of the prediction at NLO accuracy to that from the LHEs, already studied in ref. [13], where a fairly uniform increase of the distributions was found from the LHEs. Here we find that the effect of the PS is to 'bring back' the predictions close to the NLO ones with two exceptions: (i) the rapidity distribution of the hardest b-jet, which remains close to the prediction from LHEs even after shower evolution, and (ii) the distributions of transverse momenta for small $p_\perp$ (below about $50\,\text{GeV}$). Thus, the effect of the PS is in general small on top of the predictions at NLO level when the t-quarks are kept stable, or looking from an experimental point of view, when those are reconstructed from their decay products.

Secondly we study the effect of the full SMC by comparing differential distributions from the LHEs, after decay and after SMC, corresponding to the full simulation including PS, hadronization and hadron decay. The corresponding predictions are labelled as 'decay' and 'SMC' and the ratio of these is also shown in the lower panel of each plot. Here we find significant changes in almost all distributions. The decay of the heavy quarks modifies the shapes, which are further modified by the SMC in the case of distributions of transverse momenta, or invariant mass of the $b\bar{b}$-jet pair. In case of rapidity distributions the effect of the SMC is negligible.
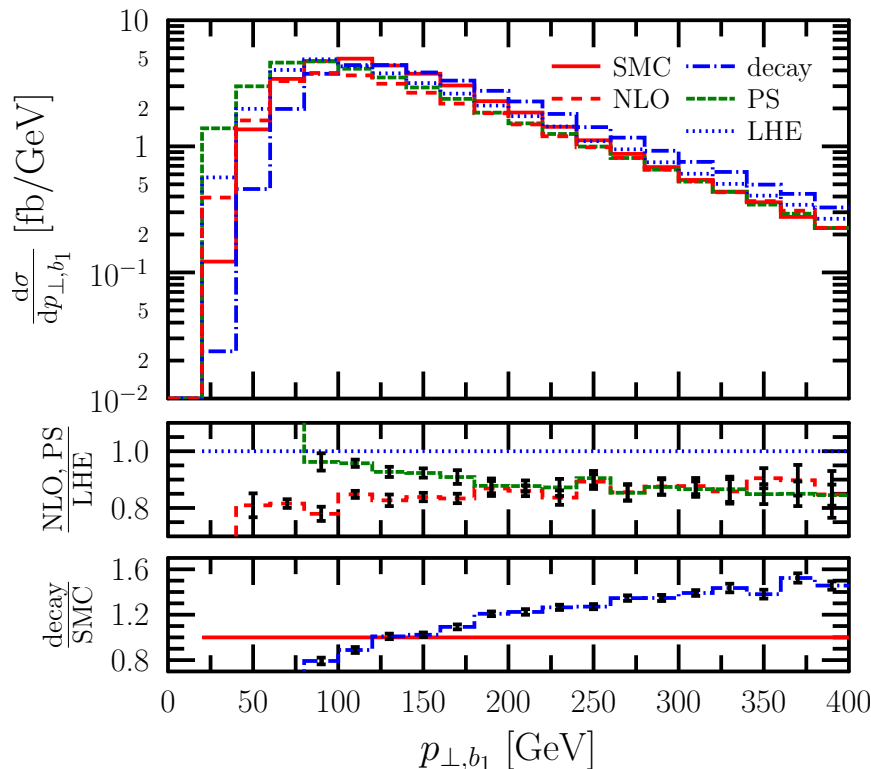
Figure 1. Distribution of transverse momentum of the hardest b-jet at the LHC at $\sqrt{s} = 14\,\text{TeV}$ using `PowHel`. Distributions from LHEs are denoted by LHE, while those at NLO accuracy by NLO. The middle panel shows the predictions at NLO accuracy, as well as after PS, normalized by the predictions from LHEs. The lower panel shows the predictions after decay normalized by the predictions after full SMC. The errorbars in both plots represent the combined statistical accuracy of the numerical integrations.

We select isolated leptons with $p_{\perp,\ell} > 25\,\text{GeV}$, $|\eta_\ell| < 2.5$ and minimal separation from all jets in the pseudorapidity-azimuthal angle plane $\Delta R = 0.4$. Looking at kinematic distributions of those leptons, such as the transverse momentum and pseudorapidity of the hardest isolated lepton plotted in figure 7, it is clear that these are not affected by SMC effects, as expected. Such leptons appear in the decay of the heavy quarks, and neither the PS nor the hadronization change those. The same is true for the W transverse mass, reconstructed for W's decaying leptonically, not shown here.

On the other hand, the distributions most affected by the SMC are the total number of $(b + \bar{b})$-jets, the total number of jets (b and non-b), the $p_\perp$ of the hardest non-b jet. We show the last of these in figure 8. In this case the prediction at NLO accuracy diverges for small transverse momentum, which is screened by the Sudakov factor, in case of parton shower matched prediction. The decay of the heavy quarks hardens this distribution significantly through jets emerging from the hadronic decay of the W bosons, which is subsequently softened by the PS and hadronization.
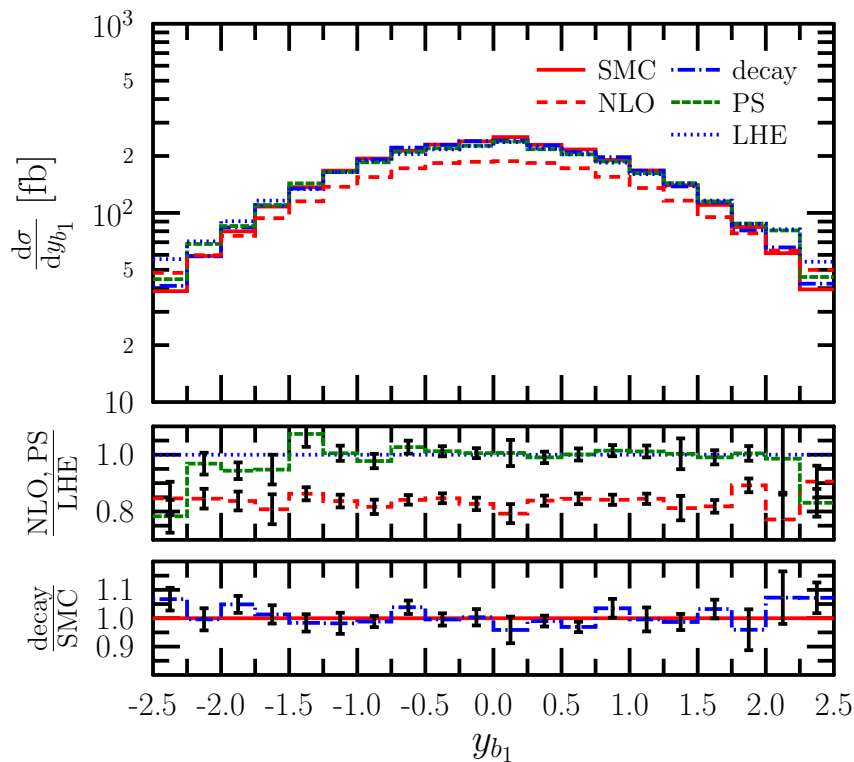
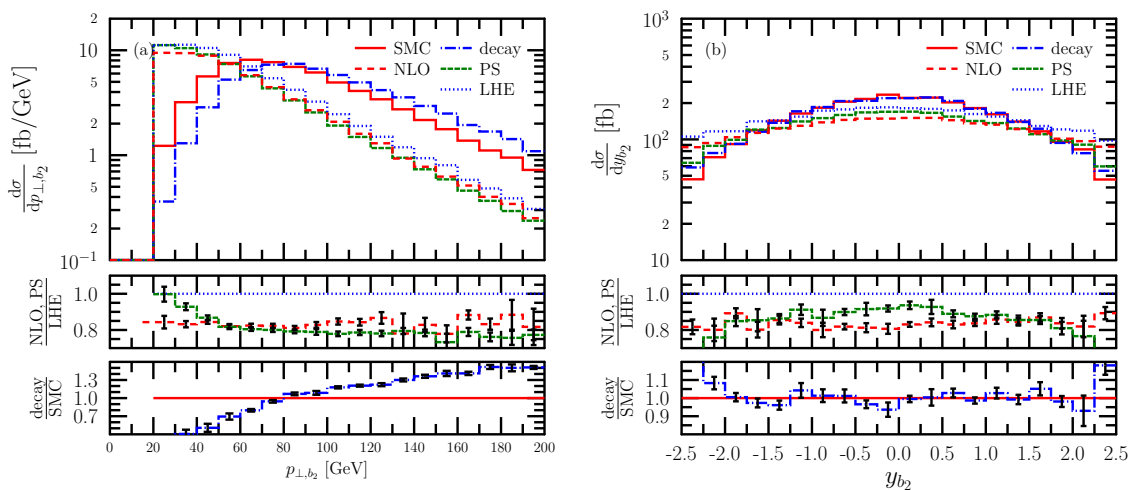**Figure 2**. Same as figure 1, but for the rapidity distribution of the hardest b-jet.

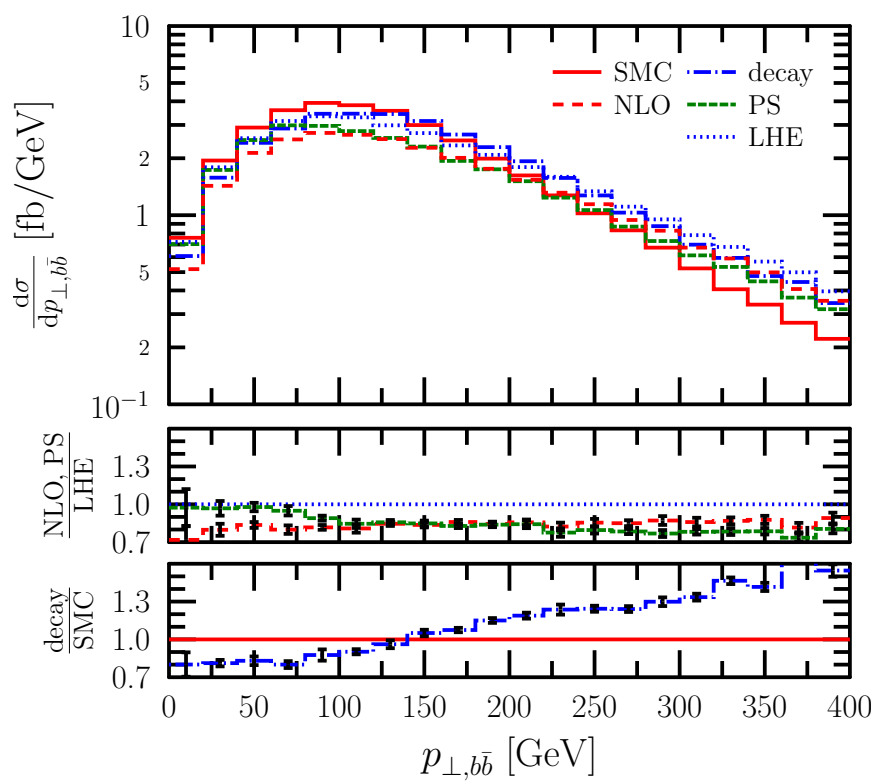**Figure 3**. Same as figures 1 and 2, but for the second hardest b-jet.

Figure 4. Same as figure 1, but for the distribution of transverse momentum of the $b\bar{b}$-jet pair.
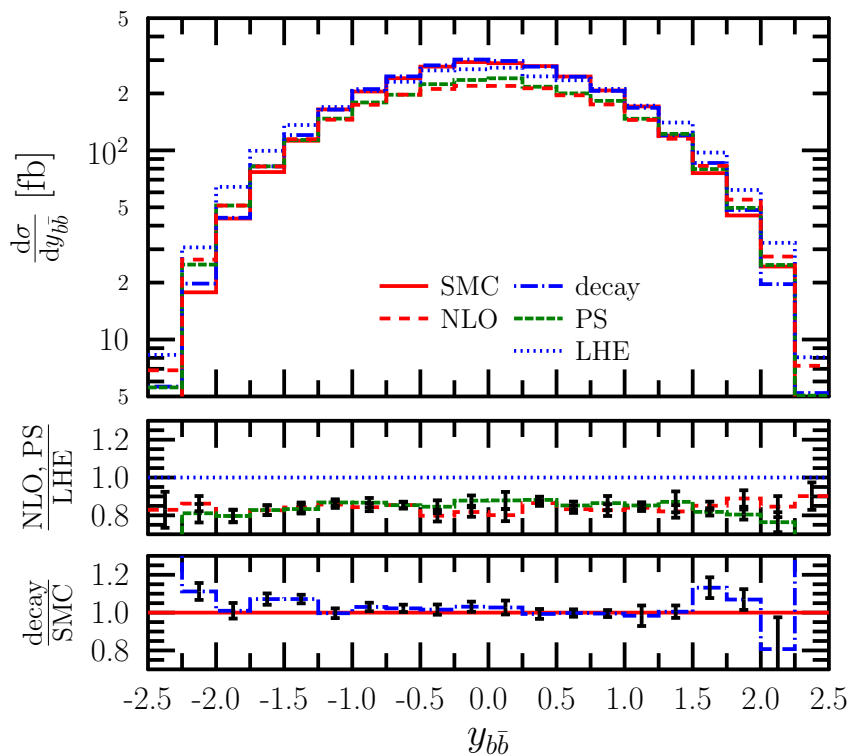
Figure 5. Same as figure 1, but for the rapidity distribution of the hardest $b\bar{b}$-jet pair.
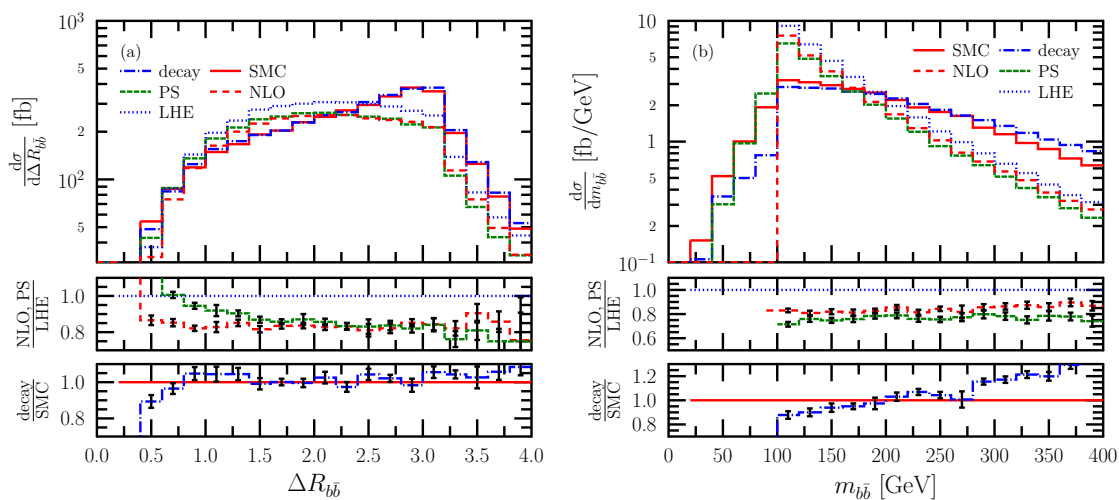

Figure 6. Same as figure 1, but for the distribution of (a) $\Delta R_{b\bar{b}}$ separation and (b) invariant mass of the hardest $b\bar{b}$-jet pair.
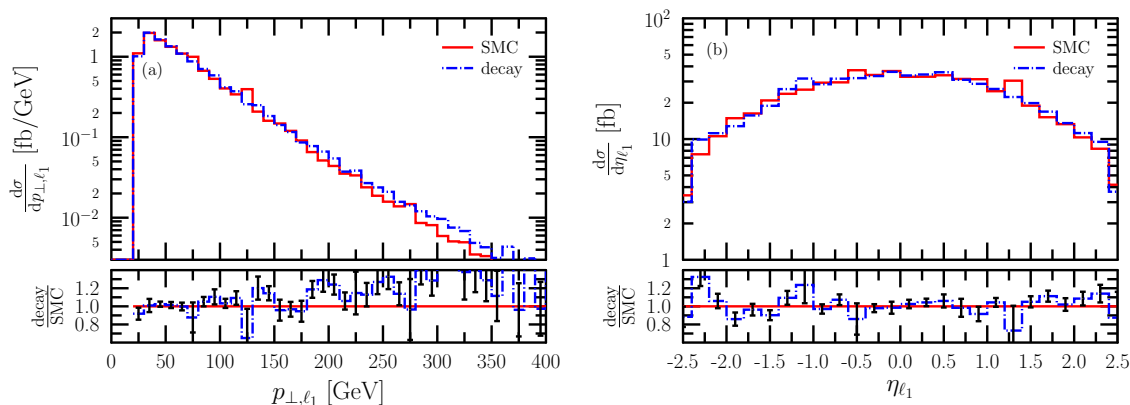
Figure 7. PowHel+PYTHIA predictions after decay and after SMC for (a) the transverse momentum and (b) the pseudorapidity of the hardest lepton. The lower panel shows the ratio of the predictions after SMC to those after decay.
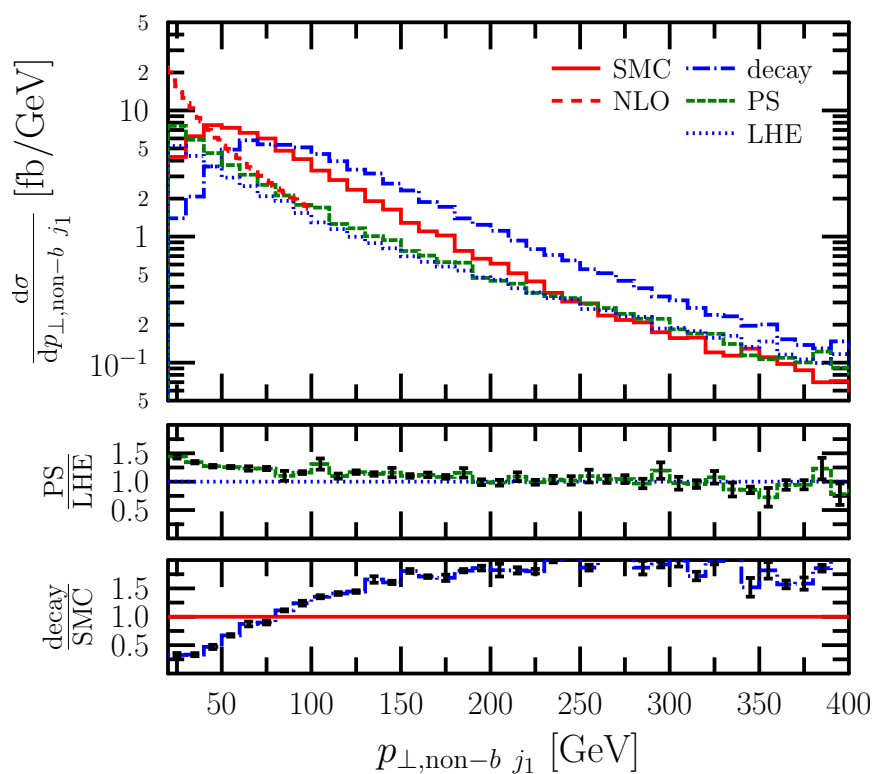

Figure 8. Same as figure 1, but for the distribution of the transverse momentum of the hardest non-b jet.

3.2 Analyses with cuts at the hadron level

In order to make predictions at the hadron level including scale and PDF uncertainties, we decided to use selection cuts at the hadron level, inspired by the CMS note ref. [35], to identify events in the dileptonic decay mode. These are rather involved, which amply shows the flexibility of our method:

1. We require the events to have at least one pair of isolated opposite sign leptons with $p_{\perp,\ell} > 20\,\text{GeV}$ and $|\eta_\ell| < 2.4$. Dilepton candidate events with invariant mass $m_{\ell\ell'} < 12\,\text{GeV}$ are removed to suppress multijet final states, at essentially no penalty for the collected signal. A lepton is considered isolated if the sum of the transverse momenta of the charged hadrons, neutral hadrons and photons in a cone of $\Delta R < 0.3$ around it divided by its transverse momentum, I_{rel}, does not exceed $I_{\text{rel}}^{\max} = 0.15$.

2. To remove the large background from events with Z-boson and jets, with Z decaying into leptons, we require that the invariant mass of the lepton pair defined above (e^+e^- or $\mu^+\mu^-$) falls outside a $\pm 15\,\text{GeV}$ window centered at m_Z.

3. Signal events typically contain large missing transverse energy due to the presence of neutrinos from decays of the W-bosons. We require a minimum missing transverse energy, $\not{p}_\perp > 30\,\text{GeV}$, for the e^+e^- or $\mu^+\mu^-$ dilepton final states, but not for the mixed flavour dilepton final state.

4. We cluster jets from hadrons, photons and non-isolated leptons using the anti-$k_\perp$ algorithm [36] with $R = 0.5$. The jets are all required to have $|\eta_j| < 2.5$ and $p_{\perp,j} > 40$ or $20\,\text{GeV}$. We performed the analysis with both cuts and report in the following the corresponding cross-sections, but present distributions here for the former only. The trend of the distributions related to the latter in fact do not present particular additional features which modify our general conclusions.

5. We require at least four well-separated b-jets after SMC, $\Delta R(j_i, j_k) > 0.5$ for i, $k \in \{1, 2, 3, 4\}$, as well as $\Delta R(j_i, \ell_k) > 0.5$ for $i \in \{1, 2, 3, 4\}$, and $k \in \{1, 2\}$ for the selected opposite-sign leptons. We assume 100% b-tagging efficiency using MCTRUTH.

With these set of cuts the cross sections for various choices of the default scale and for different PDF sets are presented in tables 1 and 2 for a collider energy of 14 TeV and in tables 3 and 4 for a collider energy of 8 TeV. These cross-section values were obtained at the hadron level by showering events generated by PowHel with a p_T-ordered version of the PYTHIA PS, by using the Perugia 2011 tune [37]. Using the default version of PYTHIA, where the PS is virtuality-ordered in absence of tunes, gives rise to values a few percent lower.

In table 5 we exhibit the numbers of expected events for an integrated luminosity of $19.6\,\text{fb}^{-1}$ at 8 TeV, for minimum transverse momentum of the jets fixed to both 40 and 20 GeV. These marginally agree with the numbers of events expected in the CMS data sample, obtained on the basis of a LO+PS computation by MadGraph [38] + PYTHIA (first line in table 2 and table 1 of ref. [35], respectively). However, this comparison can

μ_0	PDF	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	Total
$H_\perp/4$	CT10	2.71 ± 0.04	2.72 ± 0.04	7.81 ± 0.06	13.24 ± 0.09
$H_\perp/2$	CT10	1.97 ± 0.03	1.97 ± 0.03	5.62 ± 0.06	9.56 ± 0.08
$H_\perp$	CT10	1.43 ± 0.03	1.42 ± 0.02	4.19 ± 0.10	7.04 ± 0.11
$H_\perp/2$	NNPDF	2.15 ± 0.04	2.14 ± 0.04	6.17 ± 0.14	10.46 ± 0.15
$H_\perp/2$	MSTW	2.19 ± 0.04	2.18 ± 0.04	6.41 ± 0.14	10.78 ± 0.15

Table 1. Total cross sections in fb at $14\,\text{TeV}$ in the different dilepton channels with cuts listed in the text. The minimum transverse momentum of the jets is set to $40\,\text{GeV}$. The quoted uncertainties are of statistical origin.

μ_0	PDF	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	Total
$H_\perp/4$	CT10	10.88 ± 0.29	10.91 ± 0.33	32.43 ± 0.87	54.23 ± 0.98
$H_\perp/2$	CT10	7.82 ± 0.30	7.82 ± 0.35	22.19 ± 0.53	37.84 ± 0.70
$H_\perp$	CT10	5.41 ± 0.25	5.60 ± 0.27	15.96 ± 0.36	26.97 ± 0.52
$H_\perp/2$	NNPDF	8.46 ± 0.26	8.55 ± 0.26	24.80 ± 0.58	41.81 ± 0.69
$H_\perp/2$	MSTW	8.97 ± 0.50	8.59 ± 0.21	24.72 ± 0.80	42.28 ± 0.97

Table 2. Total cross sections in fb at $14\,\text{TeV}$ in the different dilepton channels with cuts listed in the text. The minimum transverse momentum of the jets is set to $20\,\text{GeV}$.

μ_0	PDF	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	Total
$H_\perp/4$	CT10	0.490 ± 0.007	0.485 ± 0.006	1.421 ± 0.010	2.396 ± 0.013
$H_\perp/2$	CT10	0.348 ± 0.005	0.345 ± 0.004	0.997 ± 0.006	1.690 ± 0.010
$H_\perp$	CT10	0.246 ± 0.004	0.242 ± 0.003	0.714 ± 0.008	1.204 ± 0.010
$H_\perp/2$	NNPDF	0.361 ± 0.006	0.360 ± 0.004	1.060 ± 0.013	1.781 ± 0.015
$H_\perp/2$	MSTW	0.374 ± 0.005	0.376 ± 0.004	1.098 ± 0.017	1.848 ± 0.018

Table 3. Total cross sections in fb at $8\,\text{TeV}$ in the different dilepton channels with cuts listed in the text. The minimum transverse momentum of the jets is set to $40\,\text{GeV}$.

μ_0	PDF	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	Total
$H_\perp/4$	CT10	2.242 ± 0.206	2.072 ± 0.182	6.451 ± 0.338	10.765 ± 0.436
$H_\perp/2$	CT10	1.547 ± 0.132	1.498 ± 0.086	4.482 ± 0.206	7.527 ± 0.259
$H_\perp$	CT10	1.070 ± 0.045	1.080 ± 0.064	3.176 ± 0.116	5.326 ± 0.140
$H_\perp/2$	NNPDF	1.621 ± 0.060	1.627 ± 0.062	4.707 ± 0.123	7.955 ± 0.150
$H_\perp/2$	MSTW	1.658 ± 0.084	1.709 ± 0.102	5.146 ± 0.246	8.513 ± 0.279

Table 4. Total cross sections in fb at $8\,\text{TeV}$ in the different dilepton channels with cuts listed in the text. The minimum transverse momentum of the jets is set to $20\,\text{GeV}$.

only be indicative for two reasons. On the one hand our predictions assume 100% b-tagging efficiency, while the experimental analysis has smaller. On the other hand, we require at least four b-jets, treating all b-jets on the same footing, while the experimental analysis requires only two b-jets, because they do not include the b-jets coming from t-

$p_{\perp,j}$	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	Total
$\geq 40\,\mathrm{GeV}$	$6.82^{+2.78}_{-2.00}$	$6.76^{+2.75}_{-2.02}$	$19.54^{+8.31}_{-5.56}$	$33.12^{+13.84}_{-9.58}$
$\geq 20\,\mathrm{GeV}$	$30.32^{+13.62}_{-9.35}$	$29.36^{+11.25}_{-8.19}$	$87.84^{+38.60}_{-25.59}$	$147.52^{+63.47}_{-43.13}$

Table 5. Number of expected $t\bar{t}b\bar{b}$ events at $8\,\mathrm{TeV}$ in the different dilepton channels with cuts listed in the text. The minimum transverse momentum of the jets is set to $40\,\mathrm{GeV}$ (first line) and $20\,\mathrm{GeV}$ (second line), respectively. The uncertainty corresponds to the variation of the renormalization and factorization scales around their default value by factors of half and two.
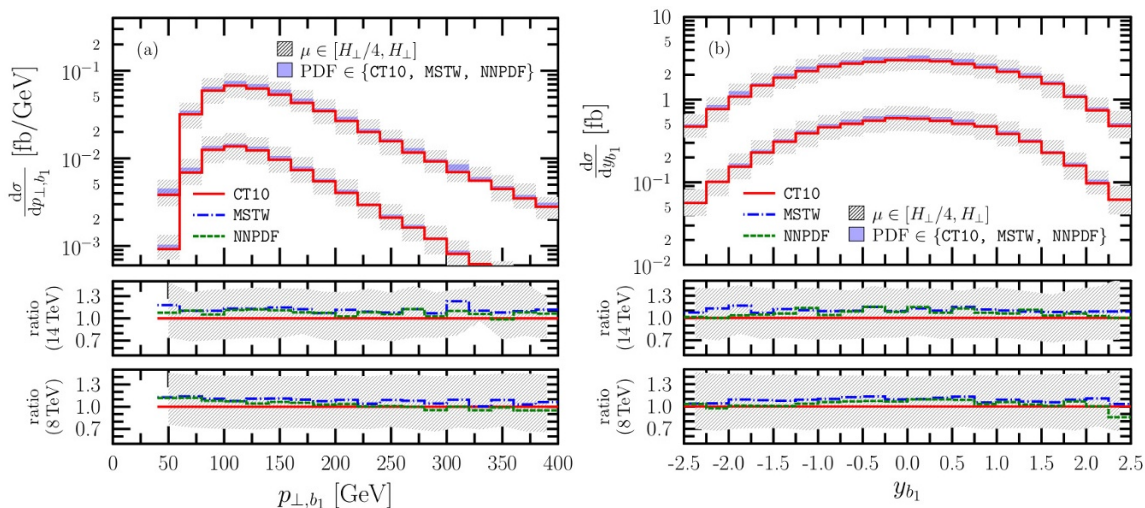
Figure 9. Distribution of (a) transverse momentum (b) rapidity of the hardest b-jet at the LHC at $\sqrt{s} = 14\,\mathrm{TeV}$ using `PowHel + PYTHIA` after full SMC, under the cuts (1–5) listed in the text. The band represent the envelope of the scale uncertainties. The lower panels show the ratio of predictions obtained with `MSTW2008NLO` and `NNPDF` sets to that obtained with `CT10NLO`.

quark decays in this count. Furthermore, in order to identify b-jets we kept the lowest lying B-hadrons stable in `PYTHIA` for simplicity, tagging as b-jets those including at least a B-hadron, whereas the experiment reconstructs b-jets from their decay products, using displaced vertex information.

Recently an independent study of $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ hadroproduction, also inspired by the same CMS note, appeared in ref. [39]. This work differs from our one because it is an NLO study, including a simplified sets of cuts applied on parton level events (also including top decays, but neglecting parton shower, hadronization and hadron decay effects).

In figures 9–12 we present several kinematic distributions obtained with these selection cuts (for the case of minimum transverse momentum of jets fixed to $40\,\mathrm{GeV}$): the transverse momenta and rapidities of the two hardest b-jets, the transverse momentum, rapidity and invariant mass distribution of the hardest b-jet pair and the invariant mass distribution of the hardest b-jet and hardest charged lepton. In each figure we show two predictions, one for $8\,\mathrm{TeV}$ c.m. energy (smaller cross sections) and the other for $14\,\mathrm{TeV}$ (larger cross sections). The shapes of the rapidity distributions are very similar at the two energies, the result of the larger collision energy just amounting to about a five to six-fold increase of the
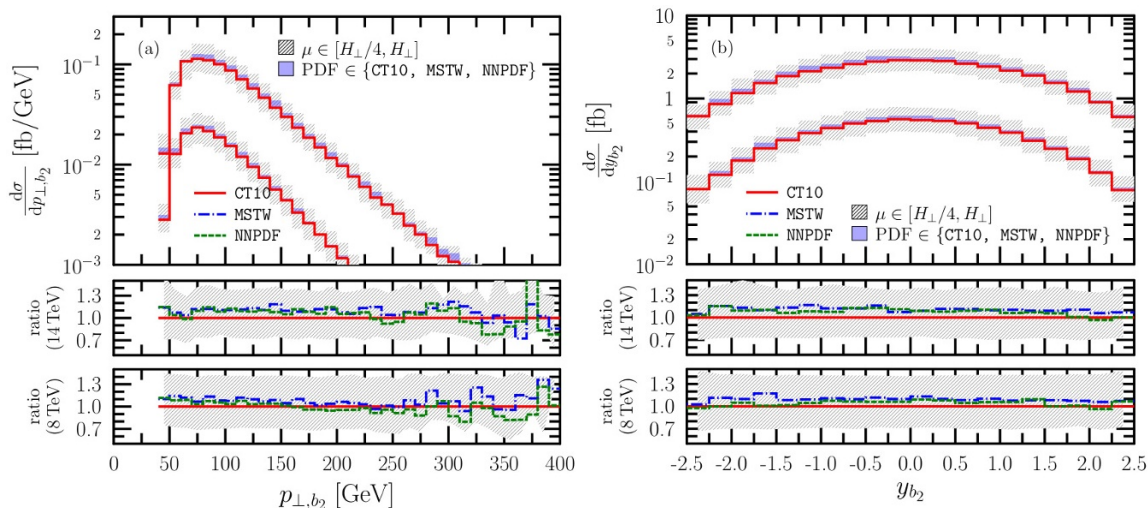
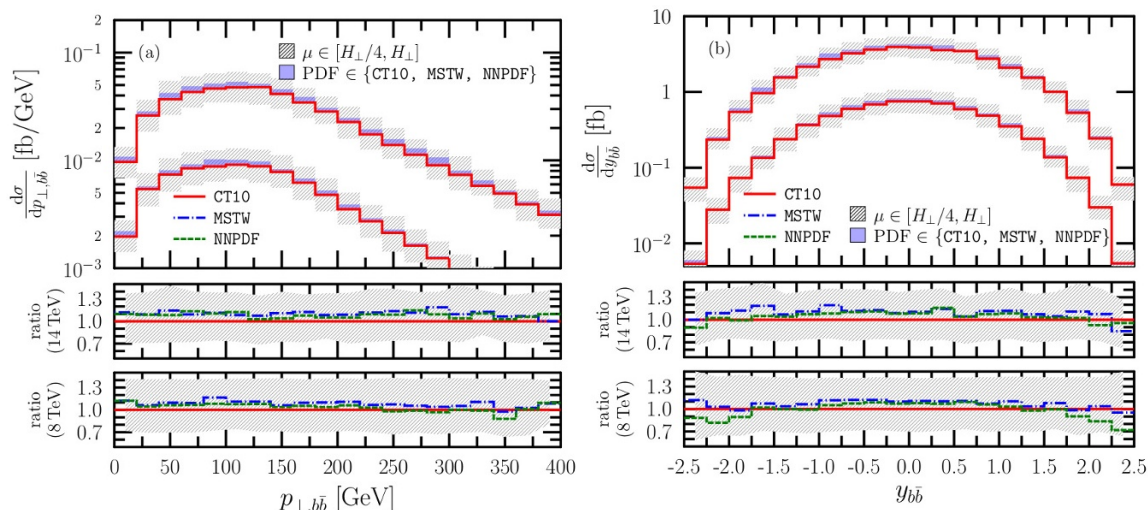**Figure 10**. Same as figure 9, as for the second hardest b-jet.

Figure 11. Same as figure 9, as for the hardest $b\bar{b}$-jet pair.

cross section. On the other hand, in case of the invariant mass and transverse momentum distributions, the spectra at 14 TeV also become considerably harder with respect to those at 8 TeV.

In the same figures we exhibit two bands: one corresponding to the envelope of standard variations of the renormalization and factorization scales, between half and twice the default scale μ_0, while the other to the envelope of dependence on the PDF set (CT10NLO, MSTW2008NLO and NNPDF). In order to see these better, in the lower panels we show all predictions normalized to our default choice: predictions with $\mu_0 = H_\perp/2$ and CT10NLO PDF set. The middle panels corresponds to the cross sections at 14 TeV, while the lower ones to those at 8 TeV.
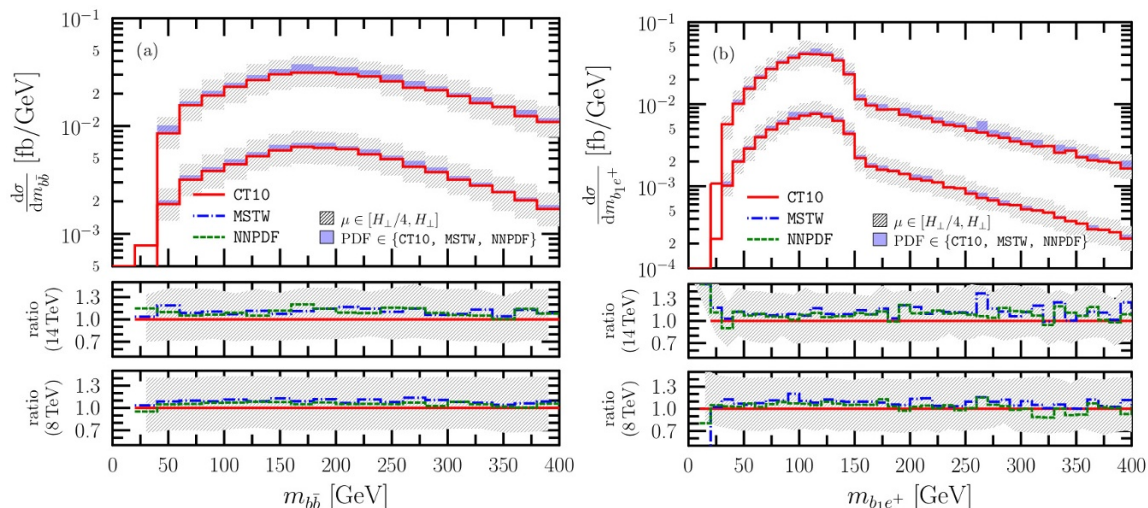
Figure 12. Same as figure 9, as for the invariant mass of (a) the hardest $b\bar{b}$-jet pair, (b) the hardest b-jet and positron.

We see that the scale dependence for these distributions follows the scale dependence found at the NLO accuracy [13]. It is similar in size (about +38%–26% at 14 TeV and +41%–29% at 8 TeV, for a jet $p_\perp$ cut at 40 GeV) and also uniform in shape.

The PDF bands are much narrower. In general, we find that up to statistical fluctuations, the smallest cross sections are obtained with `CT10NLO`, while `MSTW2008NLO` and `NNPDF` give similar, at most $\sim 10\%$ larger values than our default, with `MSTW` predictions slightly larger than `NNPDF` ones. These results show that this complicated final state with rather exclusive experimental selection cuts can be modelled at NLO accuracy matched to SMC with theoretical uncertainties that are comparable to what we can find at fixed order.

We present an example for lepton distributions in figure 13(a) where the spectrum of the transverse momentum of the hardest lepton is depicted. This lepton emerges in the decay of one of the t-quarks, i.e. it is not present in the NLO matrix elements. Nevertheless, all what we pointed out for the b-jet distributions concerning the scale and PDF uncertainties are valid for this (and other leptonic distributions, not shown here), too.

Finally, in figure 13(b) we present the distribution of the transverse momentum of the $t\bar{t}b\bar{b}$ system, emphasizing the low $p_\perp$-region, where we see the effect of Sudakov damping. This distribution is divergent for $p_\perp = 0$ at fixed order in perturbation theory, while it is finite in the matched prediction. We find decreasing scale dependence with decreasing $p_\perp$, although with naturally worsening statistics.

4 Conclusions

We have studied the $pp \to t\bar{t}b\bar{b}$ process including NLO QCD corrections computed with five massless flavours matched with parton shower, using the `PowHel` event generator interfaced with the `PYTHIA` SMC. This computation, challenging due to the presence of two massless b's in the final state, producing singular underlying Born configurations, opens the road to
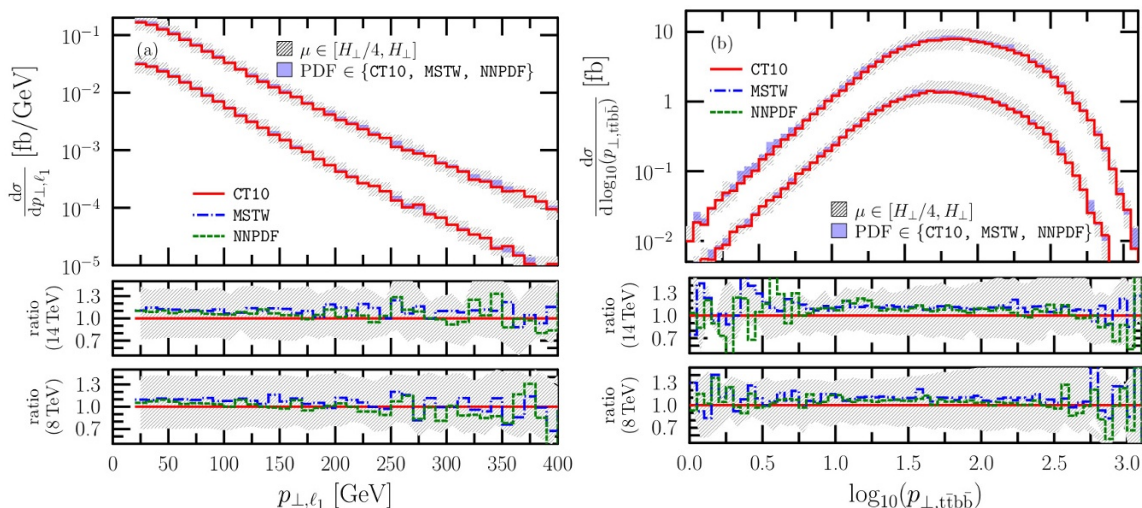
Figure 13. Same as figure 9, as for the transverse momentum of (a) the hardest lepton, (b) the $t\bar{t}b\bar{b}$ system.

a unified treatment of $t\bar{t}b\bar{b}$ and $t\bar{t}jj$. The latter process can also produce $t\bar{t}+2$ b-jet final states when both jets are tagged as b-jets due to $g \to b\bar{b}$ splittings. Such contributions are not included in our computation.

We presented predictions at different stages of event evolution in order to show the effect of the SMC in modifying NLO distributions. We found that the effect of the parton shower over the predictions at NLO accuracy are usually very small except for rapidity distributions and small values of transverse momenta. The effect of the hadronization is to soften the transverse momentum or invariant mass spectra and can be up to 30%. However, the largest effect is due to the decay of the heavy quarks. Kinematic distributions of the leptonic decay products of the heavy quarks are not affected by the hadronization.

We paid special attention to the identification of the sources of uncertainties and in the estimate of the uncertainty bands at the hadron level. We found that the scale dependences change only very moderately with c.m. energy, and are very similar to those of the predictions at the NLO accuracy reported in ref. [13], altough the latter were obtained with looser systems of cuts. These scale dependences are fairly uniform for all distributions shown when we employ our default scale, the half of the sum of transverse masses in the final state, which is a standard choice for final states with high multiplicity [28] and for those involving heavy quark pairs in multiscale dynamical regimes [29]. The PDF uncertainties are much smaller, with CT10NLO yielding the smallest cross section in general, while MSTW2008NLO and NNPDF giving very similar results that are up to about 10% higher. These results show that our events can model the $t\bar{t}b\bar{b}$ final states with uniform theoretical uncertainties similar in size to those found at fixed order.

Final states consisting of a $t\bar{t}$ and a $b\bar{b}$ quark pair constitute important backgrounds for Higgs boson production in association with a $t\bar{t}$-pair. Our events can be used to optimize the selection of the signal events to find the best signal/background ratio in the various decay channels. For this purpose sets of LHEs can be downloaded from our web-page for

both the $t\bar{t}H$ signal (http://grid.kfki.hu/twiki/bin/view/DbTheory/TthProd) and for the background (http://grid.kfki.hu/twiki/bin/view/DbTheory/TtbbProd), or new ones can be requested for different choices of scale, PDF, and parameter values.

Acknowledgments

References

[1] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, *NLO QCD corrections to* $pp \to t\bar{t}b\bar{b} + X$ *at the LHC*, *Phys. Rev. Lett.* **103** (2009) 012002 [arXiv:0905.0110] [INSPIRE].

[2] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, *NLO QCD corrections to* $t\bar{t}b\bar{b}$ *production at the LHC: 2. Full hadronic results*, *JHEP* **03** (2010) 021 [arXiv:1001.4006] [INSPIRE].

[3] G. Bevilacqua, M. Czakon, C.G. Papadopoulos, R. Pittau and M. Worek, *Assault on the NLO wishlist:* $pp \to t\bar{t}b\bar{b}$, *JHEP* **09** (2009) 109 [arXiv:0907.4723] [INSPIRE].

[4] G. Bevilacqua, M. Czakon, C.G. Papadopoulos and M. Worek, *Dominant QCD backgrounds in Higgs boson analyses at the LHC: a study of* $pp \to t\bar{t} + 2$ *jets at next-to-leading order*, *Phys. Rev. Lett.* **104** (2010) 162002 [arXiv:1002.4009] [INSPIRE].

[5] G. Cullen et al., *Next-to-leading-order QCD corrections to Higgs boson production plus three jets in gluon fusion*, *Phys. Rev. Lett.* **111** (2013) 131801 [arXiv:1307.4737] [INSPIRE].

[6] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[7] BLACKHAT collaboration, Z. Bern et al., *Universality in* $W + multijet$ *production*, [PoS(LL2014)011](PoS(LL2014)011) [arXiv:1407.6564] [INSPIRE].

[8] M. Worek, *On the next-to-leading order QCD K-factor for top* $t\bar{t}b\bar{b}$ *production at the Tevatron*, *JHEP* **02** (2012) 043 [arXiv:1112.4325] [INSPIRE].

[9] ATLAS collaboration, *Search for the standard model Higgs boson produced in association with top quarks in proton-proton collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, ATLAS-CONF-2012-135 (2012).

[10] CMS collaboration, *Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC*, *JHEP* **05** (2013) 145 [arXiv:1303.0763] [INSPIRE].

[11] CMS collaboration, *Search for Higgs boson production in association with a top-quark pair and decaying to bottom quarks or tau leptons*, CMS-PAS-HIG-13-019 (2013).

[12] CMS collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, *JHEP* **09** (2014) 087 [*Erratum ibid.* **10** (2014) 106] [arXiv:1408.1682] [CMS-HIG-13-029] [INSPIRE].

[13] A. Kardos and Z. Trócsányi, *Hadroproduction of tt̄ pair with a bb̄ pair using PowHel*, *J. Phys.* **G 41** (2014) 075005 [arXiv:1303.6291] [INSPIRE].

[14] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs cross sections: 3. Higgs properties*, arXiv:1307.1347 [INSPIRE].

[15] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244] [INSPIRE].

[16] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622] [INSPIRE].

[17] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini and F. Siegert, *NLO matching for tt̄bb̄ production with massive b-quarks*, *Phys. Lett.* **B 734** (2014) 210 [arXiv:1309.5912] [INSPIRE].

[18] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering amplitudes with open loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [arXiv:1111.5206] [INSPIRE].

[19] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146] [INSPIRE].

[20] 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] [INSPIRE].

[21] M.V. Garzelli, A. Kardos and Z. Trócsányi, *NLO event samples for the LHC*, PoS(EPS-HEP2011)282 [arXiv:1111.1446] [INSPIRE].

[22] J. Alwall et al., *A standard format for Les Houches event files*, *Comput. Phys. Commun.* **176** (2007) 300 [hep-ph/0609017] [INSPIRE].

[23] G. Bevilacqua et al., *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986 [arXiv:1110.1499] [INSPIRE].

[24] G. Bevilacqua et al., *NLO QCD calculations with HELAC-NLO*, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 211 [arXiv:1007.4918] [INSPIRE].

[25] M. Czakon, C.G. Papadopoulos and M. Worek, *Polarizing the dipoles*, *JHEP* **08** (2009) 085 [arXiv:0905.0883] [INSPIRE].

[26] 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] [INSPIRE].

[27] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [INSPIRE].

[28] CDF collaboration, T. Aaltonen et al., *Measurement of the cross section for W^- boson production in association with jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,TeV$*, *Phys. Rev.* **D 77** (2008) 011108 [arXiv:0711.4044] [INSPIRE].

[29] M. Cacciari et al., *Theoretical predictions for charm and bottom production at the LHC*, *JHEP* **10** (2012) 137 [arXiv:1205.6344] [INSPIRE].

[30] M.R. Whalley, D. Bourilkov and R.C. Group, *The Les Houches accord PDFs (LHAPDF) and LHAGLUE*, hep-ph/0508110 [INSPIRE].

[31] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, S. Dittmaier et al., *Handbook of LHC Higgs cross sections: 1. Inclusive observables*, arXiv:1101.0593 [INSPIRE].

[32] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B 406** (1993) 187 [INSPIRE].

[33] S.D. Ellis and D.E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D 48** (1993) 3160 [hep-ph/9305266] [INSPIRE].

[34] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[35] CMS collaboration, *Measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in pp collisions at $\sqrt{s} = 8\,TeV$*, CMS-PAS-TOP-13-010 (2013).

[36] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[37] P.Z. Skands, *Tuning Monte Carlo generators: the Perugia tunes*, *Phys. Rev.* **D 82** (2010) 074018 [arXiv:1005.3457] [INSPIRE].

[38] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5: going beyond*, *JHEP* **06** (2011) 128 [arXiv:1106.0522] [INSPIRE].

[39] G. Bevilacqua and M. Worek, *On the ratio of $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ cross sections at the CERN Large Hadron Collider*, *JHEP* **07** (2014) 135 [arXiv:1403.2046] [INSPIRE].