

# $t\bar{t}b\bar{b}$ at the LHC: on the size of corrections and $b$ -jet definitions

Giuseppe Bevilacqua,<sup>a</sup> Huan-Yu Bi,<sup>b</sup> Heribertus Bayu Hartanto,<sup>c</sup> Manfred Kraus,<sup>d</sup> Michele Lupattelli<sup>b</sup> and Malgorzata Worek<sup>b</sup>

<sup>a</sup>MTA-DE Particle Physics Research Group, University of Debrecen,  
H-4010 Debrecen, PBox 105, Hungary

<sup>b</sup>Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen University,  
D-52056 Aachen, Germany

<sup>c</sup>Cavendish Laboratory, University of Cambridge,  
J.J. Thomson Avenue, Cambridge CB3 0HE, U.K.

<sup>d</sup>Physics Department, Florida State University,  
Tallahassee, FL 32306-4350, U.S.A.

E-mail: [giuseppe.bevilacqua@science.unideb.hu](mailto:giuseppe.bevilacqua@science.unideb.hu),  
[bihy@physik.rwth-aachen.de](mailto:bihy@physik.rwth-aachen.de), [hbhartanto@hep.phy.cam.ac.uk](mailto:hbhartanto@hep.phy.cam.ac.uk),  
[mkraus@hep.fsu.edu](mailto:mkraus@hep.fsu.edu), [lupattelli@physik.rwth-aachen.de](mailto:lupattelli@physik.rwth-aachen.de),  
[worek@physik.rwth-aachen.de](mailto:worek@physik.rwth-aachen.de)

**ABSTRACT:** We report on the calculation of the next-to-leading order QCD corrections to the production of a  $t\bar{t}$  pair in association with two heavy-flavour jets. We concentrate on the di-lepton  $t\bar{t}$  decay channel at the LHC with  $\sqrt{s} = 13$  TeV. The computation is based on  $pp \rightarrow e^+\nu_e \mu^-\bar{\nu}_\mu b\bar{b} b\bar{b}$  matrix elements and includes all resonant and non-resonant diagrams, interferences and off-shell effects of the top quark and the  $W$  gauge boson. As it is customary for such studies, results are presented in the form of inclusive and differential fiducial cross sections. We extensively investigate the dependence of our results upon variation of renormalisation and factorisation scales and parton distribution functions in the quest for an accurate estimate of the theoretical uncertainties. We additionally study the impact of the contributions induced by the bottom-quark parton density. Results presented here are particularly relevant for measurements of  $t\bar{t}H(H \rightarrow b\bar{b})$  and the determination of the Higgs coupling to the top quark. In addition, they might be used for precise measurements of the top-quark fiducial cross sections and to investigate top-quark decay modelling at the LHC.

**KEYWORDS:** NLO Computations, QCD Phenomenology

ARXIV EPRINT: [2105.08404](https://arxiv.org/abs/2105.08404)

---

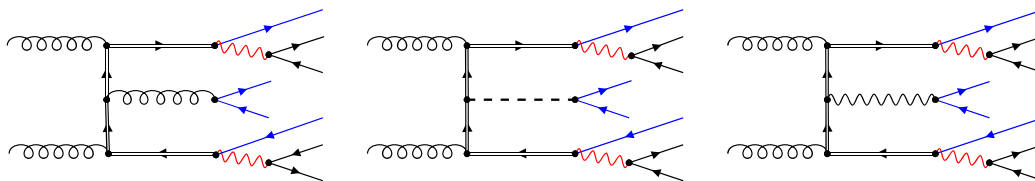
## Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b>  |
| <b>2</b> | <b>Description of the calculation and its validation</b>          | <b>4</b>  |
| <b>3</b> | <b>LHC setup</b>  | <b>7</b>  |
| <b>4</b> | <b>Integrated fiducial cross sections</b>                         | <b>10</b> |
| 4.1      | Stability test of NLO fiducial cross sections                     | 13        |
| 4.2      | Additional cuts and comparison with ATLAS results                 | 14        |
| <b>5</b> | <b>Differential fiducial cross sections and PDF uncertainties</b> | <b>16</b> |
| <b>6</b> | <b>Contribution of initial state bottom quarks</b>                | <b>23</b> |
| <b>7</b> | <b>Comparison with previous results</b>                           | <b>28</b> |
| <b>8</b> | <b>Summary</b>  | <b>29</b> |

---

## 1 Introduction

The discovery of the Higgs boson at the LHC was only the start of the wide program, the main purpose of which is to identify the properties and couplings of this recently discovered particle. In the Standard Model (SM) the Higgs boson couples to the fundamental fermions via the Yukawa interaction with a coupling strength that is proportional to the fermion mass. Probing the coupling of the Higgs boson to the top quark, the heaviest observed particle in the SM, comprises a crucial test of the consistency of the Higgs sector. Furthermore, the Top-Yukawa coupling, denoted as  $Y_t$ , might be used to constrain various models of physics beyond the SM (BSM) that very often predict a different coupling strength than the SM one. The latter is expected to be close to unity. Indirect constraints on the coupling between the top quark and the Higgs boson are available from processes including virtual top quark loops. Here the best example comprises Higgs boson production through  $gg$  fusion. On the other hand, the  $Y_t$  coupling can be probed directly in the associated production of the Higgs boson with the  $t\bar{t}$  pair, the process which has recently been observed by both the ATLAS and CMS collaborations [1, 2]. Even though  $pp \rightarrow t\bar{t}H$  contributes only about 1% to the total  $pp \rightarrow H$  production cross section, it offers a very distinctive signature. For the Higgs boson with the observed mass value the dominant decay mode is  $H \rightarrow b\bar{b}$  with the branching ratio of 58% [3]. This decay mode is additionally sensitive to the coupling of the Higgs boson to the bottom quark but it is not easily accessible experimentally. Nevertheless, both ATLAS and CMS reported searches for  $t\bar{t}H$  production in



**Figure 1.** Representative tree level Feynman diagrams for the QCD  $pp \rightarrow t\bar{t}b\bar{b}$  production process as well as for  $pp \rightarrow t\bar{t}H(H \rightarrow b\bar{b})$  and  $pp \rightarrow t\bar{t}Z(Z \rightarrow b\bar{b})$  that lead to the same  $e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b}$  final state respectively at  $\mathcal{O}(\alpha_s^4\alpha^4)$  and  $\mathcal{O}(\alpha_s^2\alpha^6)$ . The double line indicates the top (anti-top) quark, the blue line corresponds to bottom (anti-bottom) quarks, whereas the  $W$  gauge boson is depicted in red. Furthermore, the dashed line represents the SM Higgs boson.

the  $b\bar{b}$  decay channel of the Higgs boson [4, 5]. The main experimental challenge for this channel is the correct identification of the candidates for the Higgs boson decay from the so-called combinatorial background. The latter is responsible for a substantial smearing of the Higgs boson peak in the  $b\bar{b}$  invariant mass spectrum. Further challenges include the possibility of misidentification of light jets with  $b$ -jets and problems with the control of various SM backgrounds, see figure 1 for examples of Feynman diagrams for  $t\bar{t}H(H \rightarrow b\bar{b})$  production and processes that lead to the same  $t\bar{t}b\bar{b}$  final state.<sup>1</sup> With the help of  $b$ -jet tagging algorithms as well as boosted top quarks and Higgs boson it is possible to isolate the contribution of the  $t\bar{t}H(H \rightarrow b\bar{b})$  process from the most general reducible background represented by  $t\bar{t}jj$  production and from the irreducible  $Z$ -peak background [7]. Nevertheless, the  $pp \rightarrow t\bar{t}H \rightarrow t\bar{t}b\bar{b}$  process suffers greatly from the  $t\bar{t}b\bar{b}$  background, that is the most important irreducible background process for this SM Higgs boson channel. In addition, searches for four top-quark production ( $t\bar{t}t\bar{t}$ ) are also affected by the QCD  $t\bar{t}b\bar{b}$  background [8, 9]. Consequently, measurements of  $t\bar{t}H(H \rightarrow b\bar{b})$  and  $t\bar{t}t\bar{t}$  at the LHC would benefit considerably from a better understanding of the QCD  $t\bar{t}b\bar{b}$  production process and from the improved modelling of top-quark decays.

The  $pp \rightarrow t\bar{t}b\bar{b}$  process is also interesting from the theoretical point of view. This is due to the presence of two very different and distinctive scales, the top-quark mass  $m_t$  and the  $b$ -jet transverse momentum. The former governs  $t\bar{t}$  production and subsequent top-quark decays, the latter describes the two  $b$ -jets coming from the  $g \rightarrow b\bar{b}$  splitting. However, this rather straightforward picture is not adequate anymore once the contributions from off-shell top quarks and  $W$  gauge bosons are consistently taken into account. Away from the  $t\bar{t}$  threshold and for large values of  $M(b\bar{b})$  various other mechanisms start to play a non negligible part, that makes the  $t\bar{t}b\bar{b}$  process truly multi-scale production.

Calculations at NLO in QCD are available for stable top quarks already for some time [10–15]. They suffer from large uncertainties in the choice of factorisation and renormalisation scales that are of the order of 33% and large NLO corrections of the order of 77%. The latter is mainly due to the  $gg$  initial state. Indeed, NLO QCD corrections to the subprocess initiated by  $q\bar{q}$  annihilation only are small with the  $\mathcal{K}$ -factor of the order of  $\mathcal{K} = \sigma^{\text{NLO}}/\sigma^{\text{LO}} \approx 2.5\%$  whereas the NLO theoretical error due to scale variations is at

<sup>1</sup>All Feynman diagrams in this paper were produced with the help of the FEYNGAME program [6].

the 17% level [10]. Unfortunately, the large scale variation and the size of the corrections themselves for the full  $pp$  process imply that a full NNLO study would be indispensable. As the latter will remain out of reach in the nearest future, it seems that, as already suggested in ref. [13], additional kinematic restrictions to the  $b\bar{b}$  system, e.g.  $M(b\bar{b}) \gtrsim 100$  GeV or  $p_T(b\bar{b}) \gtrsim 200$  GeV as motivated by the studies of the SM Higgs boson with the mass of  $m_H = 125$  GeV, must be introduced in order to reduce large theoretical uncertainties and higher-order QCD corrections. Alternatively, a veto on extra jet radiation might be carried out to achieve the same goals. In ref. [16] even calculations of NLO QCD corrections to the  $pp \rightarrow t\bar{t}b\bar{b}j$  process were presented. For integrated NLO cross sections scale uncertainties at the level of 25% were obtained, whereas, the usual  $\mathcal{K}$ -factor was calculated to be 1.45. Having an additional resolved jet present allowed to investigate in detail the modelling of recoil effects in  $t\bar{t}b\bar{b}$  production. We note here that, generally speaking, the  $pp \rightarrow t\bar{t}b\bar{b}j$  process at NLO entails information that can be used to gain some insights into the perturbative convergence of the inclusive  $t\bar{t}b\bar{b}$  cross section beyond NLO.

Computing higher-order QCD corrections to processes with stable top quarks, no matter how technically complex they might be, can only give us a general idea of the size of the NLO corrections. Such theoretical predictions cannot provide a reliable description of the top-quark decays and are not sufficient to detail QCD radiation pattern for this process. Thus, for more realistic studies radiative top-quark decays are needed. The first step in this direction was achieved, for the  $t\bar{t}b\bar{b}$  production process, by matching the  $t\bar{t}b\bar{b}$  matrix elements, calculated with massless or massive  $b$  quarks at NLO in QCD, to parton-shower (PS) programs [17–20]. In these studies, however, top-quark decays have been either completely omitted or performed within the PYTHIA MC framework. Either way, these predictions did not include  $t\bar{t}$  spin correlations. On the other hand, in ref. [21] NLO matrix element calculations matched to a parton-shower (NLO+PS) simulations of  $pp \rightarrow t\bar{t}b\bar{b}$  were presented in the four-flavour scheme for stable top quarks and with (LO) spin correlated top-quark decays. In general, it was shown that matching and shower uncertainties were very small for this process once  $m_b \neq 0$  was considered. Moreover, neither  $\mathcal{K}$ -factor nor the size of scale uncertainties was substantially changed by PS effects. Actually, the scale dependence uncertainties even increased slightly, reaching 40%. The experimental measurements of  $pp$  production cross sections for  $t\bar{t}b\bar{b}$  have been carried out by both ATLAS and CMS [22–24]. The measured inclusive fiducial cross sections generally exceed theoretical predictions for  $t\bar{t}b\bar{b}$  as provided by the already mentioned NLO + PS simulations. Nevertheless, they are compatible within the total uncertainties. Very recently, the computation of NLO QCD corrections to  $t\bar{t}b\bar{b}$  production in the di-lepton top-quark decay channel was reported in ref. [25]. Specifically, higher order  $\alpha_s$  corrections to the  $e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b}$  final state at  $\mathcal{O}(\alpha_s^4\alpha^4)$  were calculated for the LHC energy of  $\sqrt{s} = 13$  TeV. In ref. [25] they have shown that at the level of cross section the NLO QCD corrections to  $t\bar{t}b\bar{b}$  were close to 100%. Furthermore, at the differential level on top of this overall shift, moderate shape distortions up to 25% were obtained.

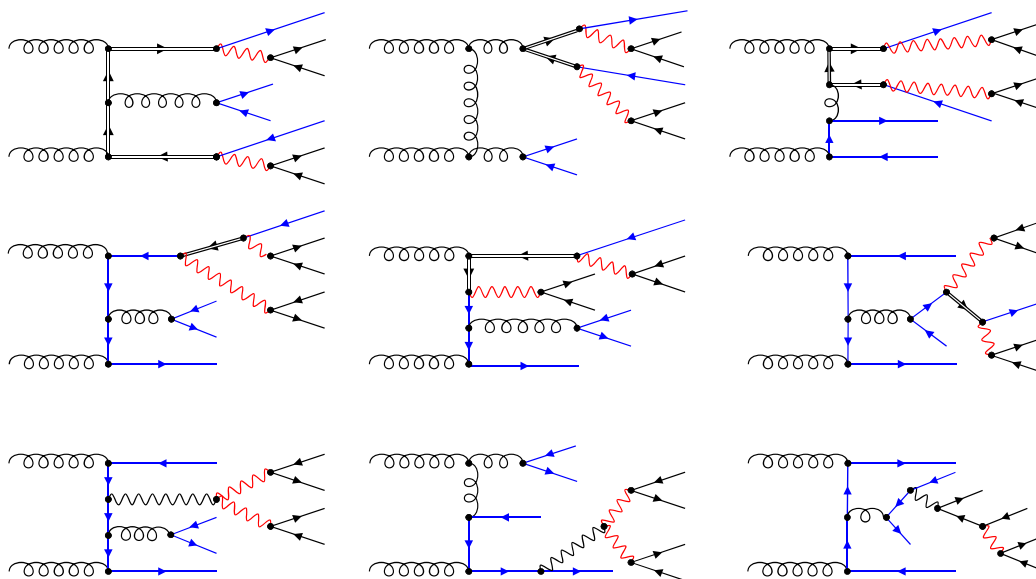
The goal of the paper is manifold. Due to the complexity of the process we will present an independent computation of the complete NLO QCD corrections to the off-shell production of  $t\bar{t}b\bar{b}$  in the di-lepton top-quark decay channel. In the first step, we

use the same scale choice and the SM input parameters as well as phase-space cuts to confirm the results presented in ref. [25] at the integrated and at the differential level. Our second goal is to extend this analysis to other selected renormalisation and factorisation scale choices and to different PDF sets. By using the error PDF sets we will additionally study the internal NLO PDF uncertainties. Furthermore, a stability test of LO and NLO fiducial cross sections with respect to the  $b$ -jet transverse momentum cut will be performed. Afterwards additional cuts will be introduced to examine their impact on the size of the  $\mathcal{K}$ -factor and the theoretical uncertainties due to scale dependence. Another goal of the paper is to investigate the initial state bottom-quark contributions and their impact on the integrated and differential  $t\bar{t}b\bar{b}$  cross sections. To this end two approaches, the so called *charge-blind* and *charge-aware* heavy-flavour jet tagging, will be introduced.

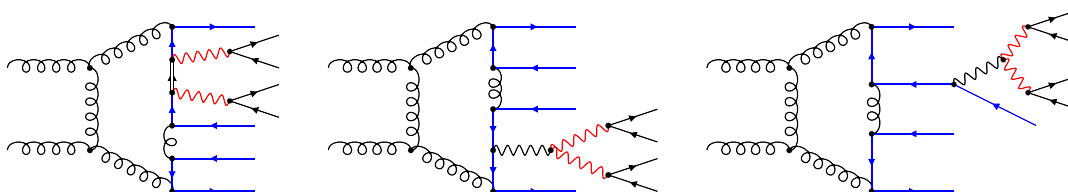
The paper is organised as follows. In section 2 we briefly describe the HELAC-NLO framework that we use for the calculation and discuss the cross-checks that have been performed. Our theoretical setup for LO and NLO QCD results is given in section 3. Phenomenological results for the integrated and differential fiducial  $t\bar{t}b\bar{b}$  cross sections are discussed in detail in section 4 and section 5. The initial state bottom-quark contribution is examined in section 6. The comparison with theoretical predictions presented in ref. [25] is performed in section 7. Finally, in section 8 our results for the  $t\bar{t}b\bar{b}$  production process are summarised.

## 2 Description of the calculation and its validation

We consider the fully realistic  $e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b} + X$  final state. We consistently take into account resonant and non-resonant top-quark contributions and all interference effects among them. In addition, non-resonant and off-shell effects due to the finite  $W$  gauge boson width are included. Due to their insignificance we neglect flavor mixing. A few examples of Feynman diagrams contributing to the leading order process at  $\mathcal{O}(\alpha_s^4\alpha^4)$  are presented in figure 2. They are shown for the dominant  $gg$  partonic subprocess. NLO QCD corrections are calculated with the help of the HELAC-NLO Monte Carlo (MC) program [26]. This is the first computation of a  $2 \rightarrow 6$  process (the decay products of the  $W$ 's are not counted, because they do not couple to colour charged states) carried out within this framework. Even though HELAC-NLO has already been employed for the calculations of NLO QCD corrections to  $t\bar{t} + X$ ,  $X = j, \gamma, Z(\rightarrow \nu\nu), W^\pm(\rightarrow \ell\nu)$  with full top quark off-shell effects included [27–31], these processes were at most  $2 \rightarrow 5$  processes from the QCD point of view. For the  $gg \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b}$  partonic reaction there are 3904 LO diagrams. For each  $q\bar{q} \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b}$  subprocess, where  $q$  stands for  $u, d, c, s$ , we have 930 LO diagrams. The calculation of the LO scattering amplitudes is performed within the HELAC-DIPOLES package [32]. The phase-space integration is performed and optimised with the help of PARNI [33] and KALEU [34]. The produced top quarks are unstable particles, thus, the inclusion of their decays is performed in the complex-mass scheme [35–38]. It fully respects gauge invariance and is straightforward to apply. The resonant electroweak vector bosons are also treated in the complex-mass scheme.



**Figure 2.** Representative tree level Feynman diagrams for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  process at  $\mathcal{O}(\alpha_s^4 \alpha^4)$ . Diagrams with two, only one and no top-quark resonances are presented. The double line indicates the top (anti-top) quark, the blue line corresponds to the bottom (anti-bottom) quark whereas the  $W$  gauge boson is depicted in red. Also shown is a diagram that contributes to the finite  $W$  width corrections.



**Figure 3.** Examples of octagon-, heptagon- and hexagon-type of one-loop diagrams contributing to the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  process. The double line indicates the top (anti-top) quark, the blue line corresponds to bottom (anti-bottom) quarks whereas the  $W$  gauge boson is depicted in red.

We compute the virtual corrections using HELAC-1LOOP [39] and CUTTOOLS [40]. Specifically, one-loop amplitudes are generated with HELAC-1LOOP and are further reduced at the integrand level using the so-called OPP method [41] as implemented in CUTTOOLS. The most complicated one-loop diagrams in our calculations are octagons (8-point integrals). In the  $gg$  channel they involve tensor integrals up to rank six. In table 1 the number of one-loop Feynman diagrams, that corresponds to each type of correction for the dominant  $gg$  partonic subprocess, is provided (see examples in figure 3). They are obtained with the help of the QGRAF program [42] as HELAC-NLO does not employ Feynman diagrams for the amplitude calculations. We have cross-checked our results with the publicly available general-purpose MC program MADGRAPH5\_AMC@NLO [43]. Specifically, we compared numerical values of the one-loop virtual corrections for a few phase-space points for  $gg$  and

| One-loop correction type | Number of Feynman diagrams |
|--------------------------|----------------------------|
| Self-energy              | 93452                      |
| Vertex                   | 88164                      |
| Box-type                 | 49000                      |
| Pentagon-type            | 25876                      |
| Hexagon-type             | 11372                      |
| Heptagon-type            | 3328                       |
| Octagon-type             | 336                        |
| Total number             | 271528                     |

**Table 1.** The number of one-loop Feynman diagrams for the dominant subprocess  $gg \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b}$  at  $\mathcal{O}(\alpha_s^5 \alpha^4)$  split by loop topology. The Higgs boson exchange contributions are not considered and the Cabibbo-Kobayashi-Maskawa mixing matrix is kept diagonal.

$q\bar{q}$  ( $q = u, d, b$ ) partonic subprocesses. Virtual corrections come from the interference of the tree-level and one-loop amplitudes, summed over all colors and spins, and for  $N_F = 5$  massless quark flavours. We compared the finite parts along with the coefficients of the poles in  $\epsilon$ . Additionally, coefficients for color and spin summed results for the  $\mathcal{I}$ -operator were cross checked between HELAC-1LOOP and HELAC-DIPOLES. We have found perfect agreement in each case. At the one loop level the appearance of a non-zero top-quark width in the propagator requires the evaluation of scalar integrals with complex masses, which is supported by the ONELOOP program [44], used for the evaluation of the integrals. The preservation of gauge symmetries by this approach is explicitly checked by studying the Ward identities up to the one-loop level. For  $gg$  subprocess we perform this test for every phase-space point. Quadruple precision is used to recompute events which fail the gauge invariance check. For the  $q\bar{q}$  subprocess we use the so-called scale test [45], which is based on momentum rescaling. Since we know how the recalculated amplitude scales when the momenta are rescaled, it is possible to compare the two results. Also in this case the test is performed for each phase-space point. For the failed points the amplitude is recomputed using higher precision. In addition, reweighting techniques, helicity and colour sampling methods are used in order to optimise the performance of the HELAC-NLO framework.

To compute the real corrections we isolate the singularities from the soft or the collinear parton emissions via subtraction methods for NLO QCD calculations. We employ HELAC-DIPOLES, which implements the dipole formalism of Catani and Seymour [46, 47] for arbitrary helicity eigenstates and colour configurations of the external partons and the Nagy-Soper subtraction scheme [48], which makes use of random polarisation and colour sampling of the external partons. Two independent subtraction schemes allow us to cross check the correctness of the real emission part of the calculation in an even more robust way. We use a phase-space restriction on the contribution of the subtraction terms as originally proposed in refs. [12, 49, 50] for the Catani-Seymour scheme and in ref. [51] for the Nagy-Soper one. We consider two extreme choices to cross check the independence

| Partonic Subprocess  | Number of Feynman diagrams | Number of CS Dipoles | Number of NS Subtractions |
|--|----------------------------|----------------------|---------------------------|
| $gg \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} g$             | 41364                      | 90                   | 18                        |
| $q\bar{q} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} g$       | 9576                       | 50                   | 10                        |
| $gq \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} q$             | 9576                       | 50                   | 10                        |
| $g\bar{q} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} \bar{q}$ | 9576                       | 50                   | 10                        |

**Table 2.** The list of partonic subprocesses contributing to the subtracted real emission at  $\mathcal{O}(\alpha_s^5 \alpha^4)$  for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} + X$  process where  $q = u, d, c, s$ . Also shown are the number of Feynman diagrams, as well as the number of Catani-Seymour and Nagy-Soper subtraction terms that correspond to these partonic subprocesses.

of the results on this parameter. All partonic subprocesses that are taken into account for the real emission contributions are listed in table 2, together with the number of the corresponding Feynman diagrams, the number of Catani-Seymour dipoles and Nagy-Soper subtraction terms. In each case, there are five times fewer terms in the Nagy-Soper subtraction scheme compared to the Catani-Seymour approach. The difference corresponds to the total number of possible spectators in the process under scrutiny, which are relevant only in the Catani-Seymour case.

Our theoretical predictions are stored in the form of modified Les Houches Event Files [52] and ROOT Ntuples [53]. Inspired by the ideas proposed in ref. [54] each “event” is stored with supplementary matrix element and PDF information. Ntuples contain unweighted events, that helps to keep the storage as small as possible. With the goal of optimising the performance of the unweighting procedure, the so-called partial unweighting is implemented in the HELAC-NLO software, see e.g. [28] for more details. Ntuples allow us to obtain theoretical predictions for different scale choices and PDF sets. Thus, for example the error PDF sets can easily be employed to calculate the internal NLO PDF uncertainties. Furthermore, any infrared-safe (IR-safe) observable can be generated, ranges and bin sizes can be adjusted while no additional time consuming running of the code is required.

### 3 LHC setup

We consider the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b} + X$  process at NLO in QCD for LHC Run II energy of  $\sqrt{s} = 13$  TeV. Specifically,  $\alpha_s$  corrections to the born-level process at  $\mathcal{O}(\alpha_s^4 \alpha^4)$  are evaluated. Different lepton generations are used to avoid virtual photon singularities stemming from  $\gamma \rightarrow \ell^+ \ell^-$ . We do not consider  $\tau$  leptons as they are usually studied separately at the LHC due to their very rich and complex decay pattern. We keep the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix diagonal. We take LO and NLO NNPDF3.1 PDF sets [55] as the default PDF sets. In both cases they are obtained with  $\alpha_s(m_Z) = 0.118$ . However, we will show results for other PDF sets, specifically for CT18 [56] and MMHT14 [57]. The difference between various PDF sets originate, among others, from the choice of the data



used and the theoretical assumptions made for the global fit. Consequently, it is desirable to see theoretical predictions also for different PDF sets. The running of the strong coupling constant  $\alpha_s$  with two-loop (one-loop) accuracy at NLO (LO) is provided by the LHAPDF interface [58] with  $N_F = 5$ . The SM input parameters for our calculations are given by

$$\begin{aligned} G_\mu &= 1.16638 \cdot 10^{-5} \text{ GeV}^{-2}, & m_t &= 173 \text{ GeV}, \\ m_W &= 80.351972 \text{ GeV}, & \Gamma_W^{\text{NLO}} &= 2.0842989 \text{ GeV}, \\ m_Z &= 91.153481 \text{ GeV}, & \Gamma_Z^{\text{NLO}} &= 2.4942664 \text{ GeV}. \end{aligned} \tag{3.1}$$

All other partons, including bottom quarks, and leptons are treated as massless particles. The top quark width is treated as a fixed parameter throughout this work and its value is evaluated for  $\alpha_s(\mu_R = m_t)$ . The top quark width at LO has been computed based on the formulas from ref. [59], whereas the NLO QCD value has been obtained upon applying the relative QCD corrections given in ref. [60] to the LO width. For our set of input parameters we have

$$\Gamma_t^{\text{LO}} = 1.443303 \text{ GeV}, \quad \Gamma_t^{\text{NLO}} = 1.3444367445 \text{ GeV}. \tag{3.2}$$

Since we treat bottom quarks as massless partons there are no diagrams with Higgs boson exchange at tree level. The Higgs boson contribution appears only at NLO through closed fermion loops involving top quarks. We checked that this contribution amounts to 0.3% of the total NLO cross section for our setup and is therefore negligible. Given its very small impact we decided to neglect the Higgs-boson contribution, which is equivalent to taking the  $m_H \rightarrow \infty$  limit. The electromagnetic coupling  $\alpha$  is calculated from the Fermi constant  $G_\mu$  ( $G_\mu$ -scheme) via

$$\alpha_{G_\mu} = \frac{\sqrt{2}}{\pi} G_F m_W^2 \sin^2 \theta_W, \quad \text{where } \sin^2 \theta = 1 - \frac{m_W^2}{m_Z^2}, \tag{3.3}$$

where  $G_\mu$  is measured in the muon decay. In the  $G_\mu$ -scheme electroweak corrections related to the running of  $\alpha$  are taken into account. The evaluation of the residual theoretical uncertainties due to the not yet calculated higher order corrections is based on the exploration of the cross section dependence on the renormalisation scale,  $\mu_R$ , and on the factorisation scale,  $\mu_F$ . For a given definition of the  $\mu_0$  scale, judiciously chosen to absorb the large logarithmic corrections that appear at higher orders, we set  $\mu_R = \mu_F = \mu_0$  but vary the two scales independently in the range

$$\frac{1}{2} \mu_0 \leq \mu_R, \mu_F \leq 2 \mu_0, \tag{3.4}$$

with the following additional condition

$$\frac{1}{2} \leq \frac{\mu_R}{\mu_F} \leq 2. \tag{3.5}$$

This means that none of the ratios  $\mu_F/\mu_0$ ,  $\mu_R/\mu_0$  and  $\mu_F/\mu_R$  can be larger than two or smaller than one-half. In this way we avoid having in the perturbative expansion logarithms

of arguments larger than a chosen amount, regardless of its arbitrariness. In practice, it comes down to considering the following pairs

$$\left(\frac{\mu_R}{\mu_0}, \frac{\mu_F}{\mu_0}\right) = \left\{ (2, 1), (0.5, 1), (1, 2), (1, 1), (1, 0.5), (2, 2), (0.5, 0.5) \right\}. \quad (3.6)$$

By searching for the minimum and maximum of the resulting cross sections one obtains an uncertainty band. The narrower the band is, the smaller the higher order corrections are expected to be. The variation of the cross section with respect to the scale choice is unphysical. It is just a reflexion of the truncation of the perturbative series. Indeed, if the cross sections are known to all orders, they will not exhibit this dependence. The scale variation is, thus, by no means a rigorous way to estimate the theoretical uncertainty. At best, it might only give an indication of the true uncertainty. For  $\mu_0$  we consider two cases. The mass of the heaviest particle appearing in the process is typically considered to be a natural scale choice. Thus, for our fixed scale setting we choose  $\mu_0 = m_t$ . Integrated fiducial cross sections are mostly influenced by final state production relatively close to the  $t\bar{t}$  threshold, which justifies our choice. However, differential cross sections extend up to energy scales that are much larger than the threshold. To this end we define the dynamical scale  $\mu_0 = H_T/3$ . The latter is given by

$$H_T = p_T(b_1) + p_T(b_2) + p_T(b_3) + p_T(b_4) + p_T(e^+) + p_T(\mu^-) + p_T^{\text{miss}}, \quad (3.7)$$

where  $b_i$ ,  $i = 1, \dots, 4$  stands for the four  $b$ -jets and  $p_T^{\text{miss}}$  is the missing transverse momentum built out of the two neutrinos ( $\nu_e, \bar{\nu}_\mu$ ). This particular choice is especially suitable for the calculations with off-shell top-quark contributions as it carries no information about the underlying top-quark resonant history. Thus, the top-quark reconstruction is not attempted here. The PDF uncertainties together with the scale dependence are the two most important ingredients of the theoretical error on the predictions of the cross sections. All final-state partons with pseudo-rapidity  $|\eta| < 5$  are recombined into jets with separation  $R = 0.4$  in the rapidity-azimuthal-angle plane via the IR-safe *anti- $k_T$*  jet algorithm [61]. Furthermore, we use the four-momentum recombination scheme. In the first part of the paper contributions induced by the bottom quarks are neglected and we require exactly four  $b$ -jets as well as two charged leptons. In the second part of the paper, when the bottom quarks are going to be included in the initial state, we will replace the requirement of having exactly four  $b$ -jets with another one, namely, we will ask for at least four  $b$ -jets in the final state. We put no restriction on the kinematics of the extra light jet (if resolved) and on the missing transverse momentum. All final state particles and  $b$ -jets have to fulfill the following criteria, which we consider to be very inclusive selection cuts

$$\begin{aligned} p_T(\ell) &> 20 \text{ GeV}, & p_T(b) &> 25 \text{ GeV}, \\ |y(\ell)| &< 2.5, & |y(b)| &< 2.5, \end{aligned} \quad (3.8)$$

where  $\ell = \mu^-, e^+$ .

## 4 Integrated fiducial cross sections

With the input parameters and cuts specified in section 3, we arrive at the following predictions for the fixed scale choice  $\mu_R = \mu_F = \mu_0 = m_t$  and the default PDF sets

$$\begin{aligned} \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = m_t) &= 6.998_{-2.569(37\%)}^{+4.525(65\%)} [\text{scales}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) &= 13.24_{-2.89(22\%)}^{+2.33(18\%)} [\text{scales}]_{-0.19(1\%)}^{+0.19(1\%)} [\text{PDF}] \text{ fb}. \end{aligned} \quad (4.1)$$

At the central scale  $\mu_0 = m_t$ , the  $gg$  channel dominates the total LO  $pp$  cross section by 94%, followed by the  $q\bar{q} + \bar{q}q$  channels with 6%. The full  $pp$  cross section receives positive and large NLO QCD corrections of 89%. The theoretical uncertainties resulting from scale variation taken in a very conservative way as a maximum of the lower and upper bounds are 65% at LO and 22% at NLO. Therefore, by going from LO to NLO we have reduced the theoretical error by a factor of 3. In the case of truly asymmetric uncertainties sometimes it is more appropriate to symmetrise the errors. After symmetrisation the scale uncertainty at LO is 51% and at NLO does not change substantially, i.e. it is reduced down to 20%. The  $\mathcal{K}$ -factor that we have obtained  $\mathcal{K} = 1.89$ , is defined as the ratio of NLO to LO cross sections. In our case both LO and NLO integrated fiducial cross sections are calculated for LO and NLO PDF sets as obtained with  $\alpha_s(m_Z) = 0.118$ . Had we used the LO NNPDF3.1 PDF set with  $\alpha_s(m_Z) = 0.130$  our LO prediction would rather be

$$\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 9.151_{-3.546(39\%)}^{+6.512(71\%)} [\text{scales}] \text{ fb}. \quad (4.2)$$

This would result in the  $\mathcal{K}$ -factor of  $\mathcal{K} = 1.45$ . Both findings are correct and reflect the different dependence on the scale of LO and NLO cross sections. Indeed, the LO cross section is much more sensitive to the variation of scales and can change more rapidly than the NLO one. Independently, we can conclude at this point that NLO QCD corrections are large and indispensable to correctly describe the process at hand.

Another source of the theoretical error comes from the parameterization of the NNPDF3.1 PDF set. These uncertainties are due to experimental errors in the various data that are used in the fits. They do not, however, take into account additional systematics coming from the underlying assumptions that enter the parametrisation of different PDF sets. The latter cannot simply be quantified within a given scheme. Therefore, we also provide NLO QCD results for two other PDF sets, namely CT18NLO and MMHT2014. We use the corresponding prescription from each group to provide the 68% confidence level (C.L.) PDF uncertainties.<sup>2</sup> Both CT18NLO PDFs and MMHT2014 PDFs include a central set as well as respectively  $N = 58$  and  $N = 50$  error sets in the Hessian representation. The NNPDF3.1 PDF set uses the MC sampling method in conjunction with neural networks. In that case the PDF uncertainties are obtained using the replicas method with a set of  $N = 100$  MC PDF members. The internal NNPDF3.1 PDF uncertainties are very small, at the level of 1% only. Our findings for CT18NLO and MMHT2014 PDF sets are, on the

<sup>2</sup>The CT18 NLO errors are rescaled since they are originally provided at 90% C.L.

other hand, given by

$$\begin{aligned}\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{CT18NLO}, \mu_0 = m_t) &= 12.85_{-2.78(22\%)}^{+2.27(18\%)} [\text{scales}]_{-0.39(3\%)}^{+0.43(3\%)} [\text{PDF}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{MMHT2014}, \mu_0 = m_t) &= 13.12_{-2.86(22\%)}^{+2.31(18\%)} [\text{scales}]_{-0.36(3\%)}^{+0.40(3\%)} [\text{PDF}] \text{ fb}.\end{aligned}\quad (4.3)$$

We can see that for CT18NLO and MMHT2014 the internal PDF uncertainties are slightly larger, i.e. of the order of 3%. These uncertainties are nevertheless similar in size to the differences between NLO QCD results obtained with various PDF sets. Indeed, the relative difference between CT18NLO and NNPDF3.1 is 3%, whereas for MMHT2014 and NNPDF3.1 we have instead 1%. We additionally present LO predictions for various LO PDF sets. Due to the lack of LO CT18 PDF sets we went back to the older version and used instead the two LO CT14 PDF sets [62]. More specifically, we employ CT14lo, CT14llo and LO MMHT2014 with  $\alpha_s(m_Z) = 0.118, 0.130$  and  $0.135$  respectively. The larger the value of  $\alpha_s(m_Z)$  the larger the resulting LO cross section would be. The latter goes to the denominator of the  $\mathcal{K}$ -factor. Our findings confirm this pattern and can be summarised as follows

$$\begin{aligned}\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{CT14lo}, \mu_0 = m_t) &= 7.098_{-2.561(36\%)}^{+4.454(63\%)} [\text{scales}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{CT14llo}, \mu_0 = m_t) &= 9.407_{-3.558(38\%)}^{+6.380(68\%)} [\text{scales}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{MMHT2014}, \mu_0 = m_t) &= 10.670_{-4.205(39\%)}^{+7.813(73\%)} [\text{scales}] \text{ fb}.\end{aligned}\quad (4.4)$$

In turn, we receive the following spread of the  $\mathcal{K}$ -factor values 1.81, 1.37, 1.23 respectively. Depending on the LO PDF set employed the NLO QCD corrections to  $pp \rightarrow t \bar{t} b \bar{b}$  production in the di-lepton top quark decay channel range from 89% to 23%. On the other hand, the NLO theoretical error is completely dominated by the scale dependence and is consistently at the 22% level (at the 20% level after symmetrisation) independently of the PDF set used.

For the dynamical scale setting  $\mu_R = \mu_F = \mu_0 = H_T/3$  our results read

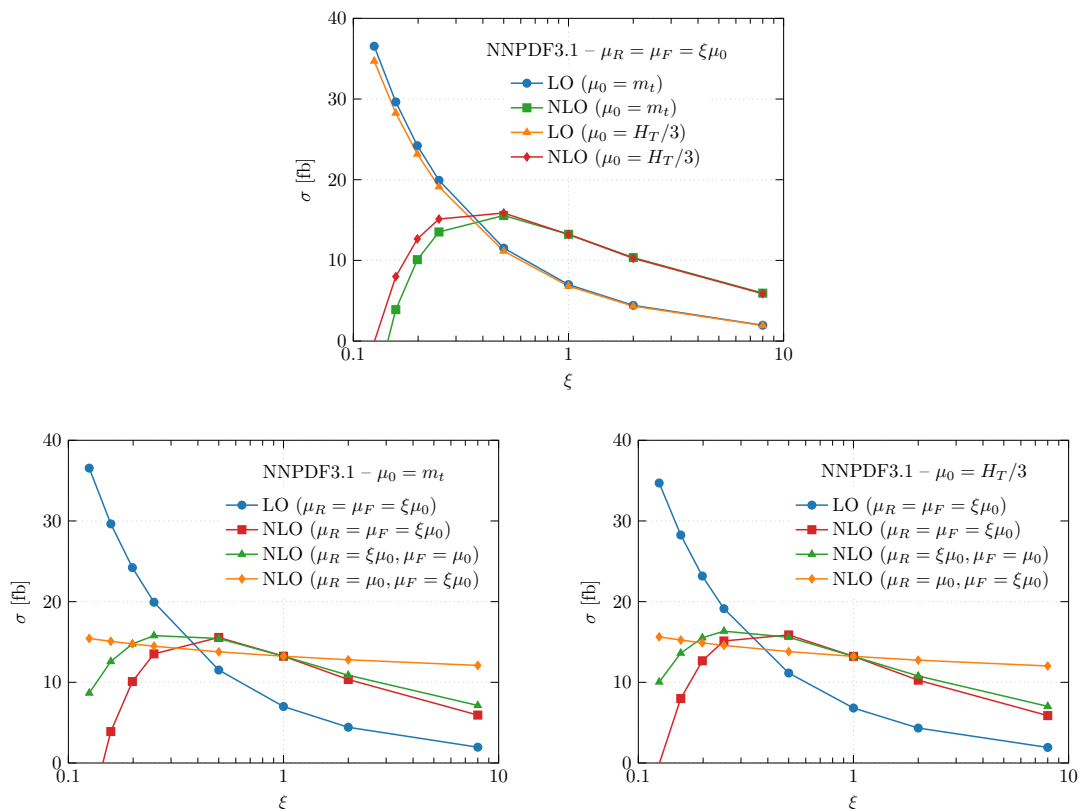
$$\begin{aligned}\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = H_T/3) &= 6.813_{-2.481(36\%)}^{+4.338(64\%)} [\text{scales}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = H_T/3) &= 13.22_{-2.95(22\%)}^{+2.66(20\%)} [\text{scales}]_{-0.19(1\%)}^{+0.19(1\%)} [\text{PDF}] \text{ fb}.\end{aligned}\quad (4.5)$$

Our NLO QCD findings for CT18NLO and MMHT2014 PDF sets are given by

$$\begin{aligned}\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{CT18NLO}, \mu_0 = H_T/3) &= 12.81_{-2.84(22\%)}^{+2.58(20\%)} [\text{scales}]_{-0.38(3\%)}^{+0.42(3\%)} [\text{PDF}] \text{ fb}, \\ \sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b}}^{\text{NLO}}(\text{MMHT2014}, \mu_0 = H_T/3) &= 13.10_{-2.92(22\%)}^{+2.64(20\%)} [\text{scales}]_{-0.35(3\%)}^{+0.40(3\%)} [\text{PDF}] \text{ fb}.\end{aligned}\quad (4.6)$$

As expected, theoretical predictions for  $\mu_0 = H_T/3$  for the integrated fiducial cross sections are very similar to the results for  $\mu_0 = m_t$ . We show these results for completeness and because we will employ the former scale choice at the differential level, where it plays a crucial role.

In figure 4, the graphical presentation of the behaviour of LO and NLO cross sections upon varying the central value of  $\mu_R$  and  $\mu_F$  by a factor of  $\xi \in \{0.125, \dots, 8\}$  is shown for the



**Figure 4.** Scale dependence of the LO and NLO integrated fiducial cross sections for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV. Renormalisation and factorisation scales are set to the common value  $\mu_R = \mu_F = \mu_0$  with  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$ . The LO and NLO NNPDF3.1 PDF sets are employed. For each case of  $\mu_0$  also shown is the variation of  $\mu_R$  with fixed  $\mu_F$  and the variation of  $\mu_F$  with fixed  $\mu_R$ .

NNPDF3.1 PDF sets. Both cases  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$  are depicted, that allowed us to compare the two scales. For the sake of completeness, in figure 4 we present again the scale dependence of the LO and NLO integrated cross sections for each case of  $\mu_0$  separately. Also shown is the variation of  $\mu_R$  with the fixed value of  $\mu_F$  and the variation of  $\mu_F$  with fixed  $\mu_R$ . We can note that from the point of view of the integrated fiducial cross sections each scale is a valid choice that might be used in phenomenological applications. We can also add that in the range  $\xi \in \{0.5, 1, 2\}$  the scale variation is driven by the changes in  $\mu_R$ . In other words, should we vary  $\mu_R$  and  $\mu_F$  up and down by a factor of 2 around  $\mu_0$  simultaneously instead of independently the scale dependence uncertainties would not change significantly. Furthermore, not only the choice of  $\mu_0$  but also the value of  $\xi$  is important. The latter if not properly selected can introduce large higher order effects and even result in negative (unphysical) NLO cross sections as can be observed in figure 4 for  $\xi \lesssim 0.25$ . Consequently, such small values should not be selected to set the scale for this process.

| $p_T(b)$                        | $\sigma^{\text{LO}}$ [fb] | $\delta_{\text{scale}}$      | $\sigma^{\text{NLO}}$ [fb] | $\delta_{\text{scale}}$    | $\delta_{\text{PDF}}$    | $\mathcal{K} = \sigma^{\text{NLO}}/\sigma^{\text{LO}}$ |
|---------------------------------|---------------------------|------------------------------|----------------------------|----------------------------|--------------------------|--|
| $\mu_R = \mu_F = \mu_0 = m_t$   |                           |                              |                            |                            |                          |  |
| 25                              | 6.998                     | +4.525 (65%)<br>-2.569 (37%) | 13.24                      | +2.33 (18%)<br>-2.89 (22%) | +0.19 (1%)<br>-0.19 (1%) | 1.89   |
| 30                              | 5.113                     | +3.343 (65%)<br>-1.889 (37%) | 9.25                       | +1.32 (14%)<br>-1.93 (21%) | +0.14 (2%)<br>-0.14 (2%) | 1.81   |
| 35                              | 3.775                     | +2.498 (66%)<br>-1.401 (37%) | 6.57                       | +0.79 (12%)<br>-1.32 (20%) | +0.10 (2%)<br>-0.10 (2%) | 1.74   |
| 40                              | 2.805                     | +1.867 (67%)<br>-1.051 (37%) | 4.70                       | +0.46 (10%)<br>-0.91 (19%) | +0.08 (2%)<br>-0.08 (2%) | 1.68   |
| $\mu_R = \mu_F = \mu_0 = H_T/3$ |                           |                              |                            |                            |                          |  |
| 25                              | 6.813                     | +4.338 (64%)<br>-2.481 (36%) | 13.22                      | +2.66 (20%)<br>-2.95 (22%) | +0.19 (1%)<br>-0.19 (1%) | 1.94   |
| 30                              | 4.809                     | +3.062 (64%)<br>-1.756 (37%) | 9.09                       | +1.66 (18%)<br>-1.98 (22%) | +0.16 (2%)<br>-0.16 (2%) | 1.89   |
| 35                              | 3.431                     | +2.191 (64%)<br>-1.256 (37%) | 6.37                       | +1.07 (17%)<br>-1.36 (21%) | +0.11 (2%)<br>-0.11 (2%) | 1.86   |
| 40                              | 2.464                     | +1.582 (64%)<br>-0.901 (37%) | 4.51                       | +0.72 (16%)<br>-0.95 (21%) | +0.09 (2%)<br>-0.09 (2%) | 1.83   |

**Table 3.** LO and NLO integrated fiducial cross sections for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b}b\bar{b} + X$  process at the LHC with  $\sqrt{s} = 13$  TeV. Results are evaluated using  $\mu_R = \mu_F = \mu_0$  with  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$ . LO and NLO NNPDF3.1 PDF sets are used. We display results for four different values of the  $p_T(b)$  cut. Also given are the theoretical uncertainties coming from scale variation ( $\delta_{\text{scale}}$ ) and PDFs ( $\delta_{\text{PDF}}$ ). In the last column the  $\mathcal{K}$ -factor is shown.

#### 4.1 Stability test of NLO fiducial cross sections

In table 3 we show the integrated fiducial cross sections at LO and NLO for different cuts on the transverse momentum of the  $b$ -jet. Theoretical uncertainties coming from scale variation are denoted as  $\delta_{\text{scale}}$ , whereas the internal NNPDF3.1 PDF uncertainties are labeled as  $\delta_{\text{PDF}}$ . Also shown is the  $\mathcal{K}$ -factor. Results are evaluated using  $\mu_R = \mu_F = \mu_0$  with  $\mu_0 = m_t$  and for the LO and NLO NNPDF3.1 PDF sets. We observe a very stable behaviour of systematics when varying the  $p_T(b)$  cut, within the 25 – 40 GeV range. Specifically,  $\delta_{\text{scale}}$  is consistently of the order of 20% and the PDF uncertainties are small for each value of the  $p_T(b)$  cut. The size of NLO QCD corrections is reduced from 89% to 68%. This reduction of 21% is well within the NLO uncertainties for this process. Results are not changed when they are generated for the dynamical scale choice. Theoretical predictions at LO and NLO for  $\mu_0 = H_T/3$  are also given in table 3. For the nominal value of the  $p_T(b)$  cut the  $\mathcal{K}$ -factor is slightly larger, i.e.  $\mathcal{K} = 1.94$ . It is reduced down to  $\mathcal{K} = 1.83$  for  $p_T(b) > 40$  GeV. The 11% difference is again well within  $\delta_{\text{scale}}$ . We can conclude that the perturbative expansion in  $\alpha_s$  for the process at hand is not spoiled by the appearance of large logarithms, thus, under excellent theoretical control.

## 4.2 Additional cuts and comparison with ATLAS results

In the following we will examine the behaviour of the integrated cross section upon adding additional cuts, i.e. making the available phase space for the  $2 \rightarrow 8$  process more exclusive. We will show results for  $\mu_R = \mu_F = \mu_0 = m_t$  only. We have checked, however, that the LO and NLO theoretical predictions for  $\mu_0 = H_T/3$  are very similar. In the first step we increase  $p_T(\ell)$  from 20 GeV to 25 GeV. Furthermore, we introduce an extra cut, i.e. the separation between the  $b$ -jet and the charged lepton in the rapidity-azimuthal-angle plane of  $\Delta R(\ell b) > 0.4$ . The NLO results with these cuts read

$$\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 9.70^{+1.66(17\%)}_{-2.10(22\%)} \text{ fb}. \quad (4.7)$$

The corresponding  $\mathcal{K}$ -factor is  $\mathcal{K} = 1.88$ . In the next step another cut has been added, i.e. the separation between charged leptons of  $\Delta R(\ell\ell) > 0.4$ . We report the following NLO cross section for this case

$$\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 9.52^{+1.62(17\%)}_{-2.06(22\%)} \text{ fb}. \quad (4.8)$$

Finally, we put the restriction on the missing transverse momentum of  $p_T^{\text{miss}} > 50$  GeV. In this case we report

$$\sigma_{pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu \bar{b} \bar{b} \bar{b}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 6.72^{+1.14(17\%)}_{-1.46(22\%)} \text{ fb}. \quad (4.9)$$

These two additional conditions for  $\Delta R(\ell\ell)$  and  $p_T^{\text{miss}}$  do not affect the  $\mathcal{K}$ -factor. It remains at the  $\mathcal{K} = 1.88$  level. We also notice that the theoretical uncertainties due to scale dependence are remarkably stable for all three cases. A few comments are in order. The first selection that we have applied, namely

$$\begin{aligned} p_T(\ell) &> 25 \text{ GeV}, & p_T(b) &> 25 \text{ GeV}, \\ |y(\ell)| &< 2.5, & |y(b)| &< 2.5, \\ \Delta R(bb) &> 0.4, & \Delta R(\ell b) &> 0.4, \end{aligned} \quad (4.10)$$

with the *anti* -  $k_T$  jet algorithm and the radius parameter  $R = 0.4$ , is very close to the phase-space volume in which the fiducial  $t\bar{t}b\bar{b}$  cross section has recently been measured by the ATLAS collaboration in the  $e\mu$  top-quark decay channel [23]. In that paper, among others,  $\sigma_{t\bar{t}b\bar{b}}$  was determined by requiring exactly one electron and one muon (with opposite charges) and at least four  $b$ -jets. We shall label this final state as  $e\mu + 4b$  in the following. The measured cross section, after the estimated contributions from  $t\bar{t}H, t\bar{t}W^\pm$  and  $t\bar{t}Z$  productions were subtracted, was compared with the theoretical predictions for the  $t\bar{t}b\bar{b}$  process. The experimental result was found to be higher than the SM prediction but still compatible within the quoted uncertainties. In detail, the measured ATLAS result is given by

$$\sigma_{e\mu+4b}^{\text{ATLAS}} = (25 \pm 6.5) \text{ fb}, \quad (4.11)$$

where 6.5 fb (26%) is the total uncertainty for this measurement, i.e. statistical plus systematical uncertainty. In table 4 we show various theoretical predictions for  $t\bar{t}b\bar{b}$  production that have been used in ref. [23] to compare with the ATLAS data. In particular,

results shown there have been generated with the following MC frameworks: SHERPA+OPENLOOPS [18], POWHEL+PYTHIA 8 [17, 19, 20] and MADGRAPH5-AMCNLO+POWHEG-BOX+PYTHIA 8 [43], that are commonly used by the ATLAS and CMS experimental collaborations. In all cases massive  $b$ -quarks and four-flavour PDF sets have been employed. The second POWHEL result has been obtained for massless  $b$ -quarks with the five-flavour PDF set. All theoretical results are based on the NLO matrix element calculations for the on-shell  $t\bar{t}b\bar{b}$  process. Let us note here, that the measured inclusive fiducial cross section exceed theoretical predictions for the  $t\bar{t}b\bar{b}$  process. However, they are all compatible within the given uncertainties. For comparison, we also add our result multiplied by a factor of 2 to account for the two decay channels  $e^+\mu^-$  and  $e^-\mu^+$ . We can observe that the  $\sigma_{e\mu+4b}^{\text{HELAC-NLO}(5\text{FS})}$  result is closer to the  $\sigma_{e\mu+4b}^{\text{ATLAS}}$  one than other theoretical predictions given in table 4. Both results agree very well within the quoted uncertainties. Specifically, the agreement of  $0.7\sigma$  has been obtained.<sup>3</sup> A dedicated comparison between HELAC-NLO and predictions obtained with the help of various NLO matrix element calculations matched to parton shower programs would be in order to understand the source of the spread among these theoretical results. We leave such studies for the near future. We note that in the experimental result as well as in the theoretical predictions used in ref. [23] also leptonic  $\tau^\pm$  decays were included, i.e.  $e^\pm$  and  $\mu^\pm$  from  $\tau^\pm$  decays were incorporated into the analysis. This is not the case for the HELAC-NLO simulations. Nonetheless, we can roughly estimate the size of the missing contribution by using the NLO fiducial cross section for  $pp \rightarrow \tau^+\nu_\tau \tau^-\bar{\nu}_\tau b\bar{b}b\bar{b} + X$  multiplied by the corresponding branching ratio for the leptonic  $\tau^+\tau^-$  decays. Since our selection cuts are very inclusive this estimate should not be very far away from the true result. Assuming the following branching ratios  $\mathcal{BR}(\tau^- \rightarrow \mu^-\bar{\nu}_\mu\nu_\tau) = 17.39\%$  and  $\mathcal{BR}(\tau^- \rightarrow e^-\bar{\nu}_e\nu_\tau) = 17.82\%$  [63] we can estimate that the HELAC-NLO result should be increased by about 0.6 fb. Thus, the final theoretical result is rather given by

$$\sigma_{e\mu+4b}^{\text{HELAC-NLO}} = (20.0 \pm 4.3) \text{ fb}, \tag{4.12}$$

which is only  $0.6\sigma$  away from the ATLAS measured cross section. We conclude this part by saying that the precision of the measurement is still slightly lower than that of the theoretical prediction as obtained with the help of the HELAC-NLO MC program. In the former case the experimental total error is at the 26% level. It comprises statistical and systematical uncertainties. In the latter case the estimated theoretical error is 22%. It combines scale dependence and PDF uncertainties.

---

<sup>3</sup>We note that the scale choice used in ref. [23] is different than our dynamical scale setting  $\mu_R = \mu_F = \mu_0 = H_T/3$ . The four theoretical predictions from ref. [23] employ  $\mu_R = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$ , where  $E_{T,i}$  refers to the transverse energy. The factorisation scale is set to  $\mu_F = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b},j} E_{T,i}$ . Due to the stable top quarks, that appear in the definition of  $\mu_R$  and  $\mu_F$ , we cannot use the exact same scale settings. We can, however, mimic it as close as possible by using the reconstructed top-quark momenta. Had we used this approach our NLO QCD results would rather be  $\sigma_{e\mu+4b}^{\text{HELAC-NLO}} = (20.3 \pm 4.2) \text{ fb}$  instead of  $\sigma_{e\mu+4b}^{\text{HELAC-NLO}} = (19.4 \pm 4.2) \text{ fb}$ .



| Theoretical predictions   | $\sigma_{e\mu+4b}$ [fb] |
|---------------------------|-------------------------|
| SHERPA+OPENLOOPS (4FS)    | $17.2 \pm 4.2$          |
| POWHEG-BOX+PYTHIA 8 (4FS) | 16.5                    |
| POWHEL+PYTHIA 8 (5FS)     | 18.7                    |
| POWHEL+PYTHIA 8 (4FS)     | 18.2                    |
| HELAC-NLO (5FS)           | $19.4 \pm 4.2$          |

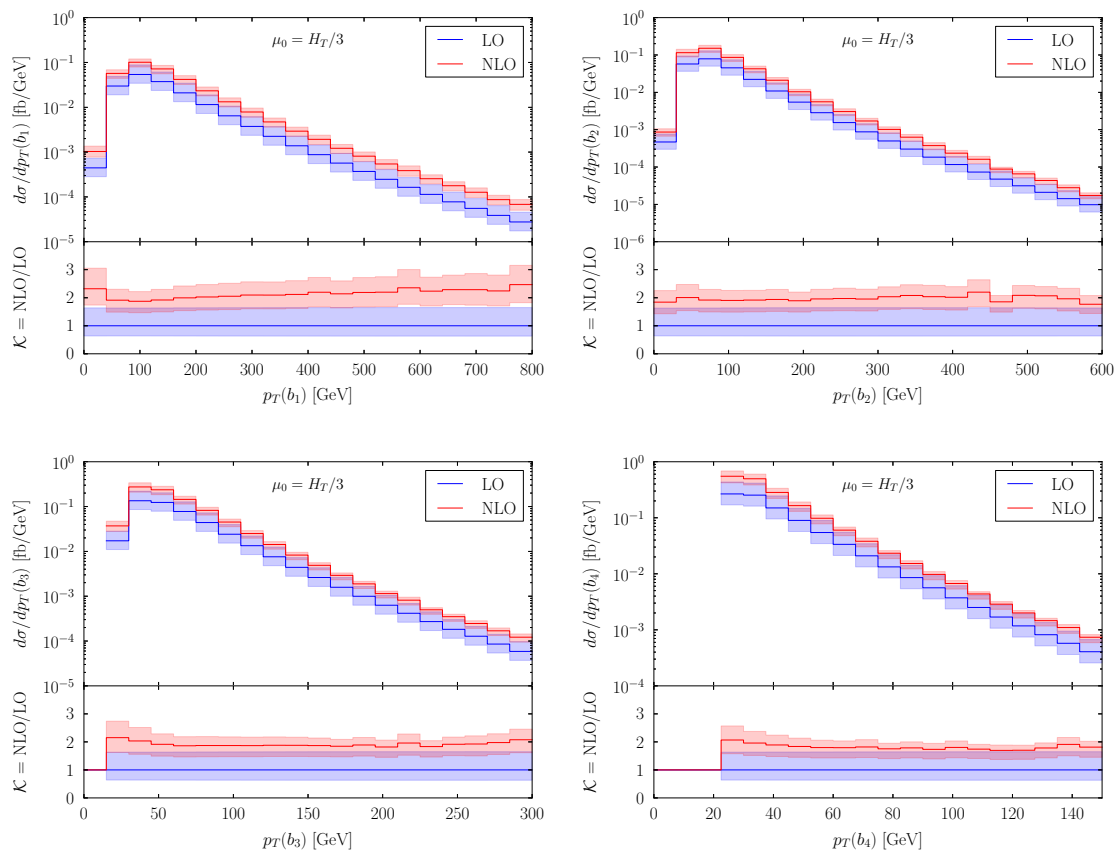
**Table 4.** Predicted fiducial cross section results for the  $pp \rightarrow t\bar{t}b\bar{b}$  process in the  $e\mu$  top quark decay channel at the LHC with  $\sqrt{s} = 13$  TeV. Results for the case of at least four  $b$ -jets in the final state are shown. Except from the HELAC-NLO case all results are taken from ref. [23].

## 5 Differential fiducial cross sections and PDF uncertainties

In addition to the normalization of the integrated fiducial cross section higher-order QCD corrections can affect the shape of various kinematic distributions. To quantify the size of these effects and to investigate shape distortions that are introduced in this way we shall test a variety of differential distributions for the  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b} + X$  process. All observables are obtained for the  $\mu_R = \mu_F = \mu_0 = H_T/3$  scale choice, NNPDF3.1 PDF sets and default cuts and parameters. Even though we present results for the kinematic-dependent choice of  $\mu_R$  and  $\mu_F$  only we have also studied differential predictions for the kinematic-independent scale choice, i.e. for  $\mu_R = \mu_F = \mu_0 = m_t$ . We will use the latter findings to provide relevant comments. However, in order not to lengthen the manuscript unnecessarily, they will not be shown. This is additionally justified by the fact that at the differential level  $\mu_0 = m_t$  cannot describe efficiently the multi-scale kinematics of the  $t\bar{t}b\bar{b}$  process. The fixed scale setting leads to perturbative instabilities in the TeV region, where large negative corrections are visible in the tails of several distributions, see e.g. ref. [28] for a detailed discussion of this behaviour. We show in the upper plots the absolute LO and NLO QCD predictions for the  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b} + X$  process. The lower panels display the differential  $\mathcal{K}$ -factors defined as  $\mathcal{K} = d\sigma^{\text{NLO}}(\mu_0)/d\sigma^{\text{LO}}(\mu_0)$ . Additionally, we provide the uncertainty bands from scale variation defined according to  $\mathcal{K}(\mu = \xi\mu_0) = d\sigma^{\text{NLO}}(\mu = \xi\mu_0)/d\sigma^{\text{LO}}(\mu_0)$ . The LO blue bands are given to illustrate the relative scale uncertainty of the LO cross section. They are defined according to  $d\sigma^{\text{LO}}(\mu = \xi\mu_0)/d\sigma^{\text{LO}}(\mu_0)$ . The LO and NLO uncertainty bands are obtained by performing a 7-point scale variation around the central value  $\mu_0$ .

We note here that our analysis is not trying to identify the origin of  $b$ -jets, i.e. it does not distinguish between the extra (prompt)  $b$ -jets and  $b$ -jets that come from the top-quark decays. This is to avoid the use of the reconstruction techniques for assigning  $b$ -jets to a specific production process. We leave such topic for future studies in which modeling uncertainties will be explored in more detail.

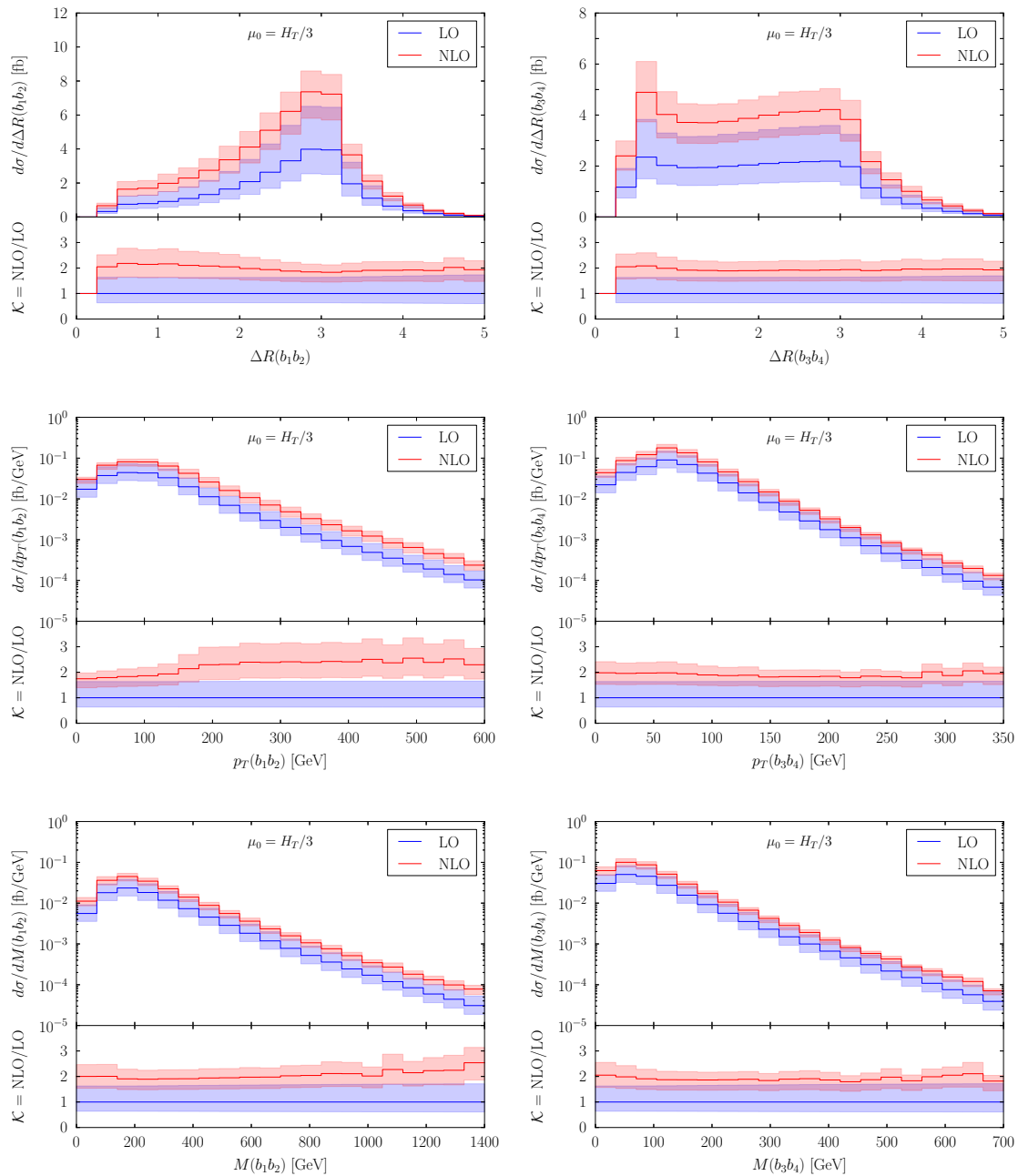
We start with the transverse momentum of the four  $b$ -jets. They are depicted in figure 5. The heavy-flavour jets are ordered in  $p_T$ , thus,  $p_T(b_1)$  corresponds to the leading (hardest)  $b$ -jet while  $p_T(b_4)$  is the fourth-leading (softest)  $b$ -jet. The differential cross sec-



**Figure 5.** Differential cross section distributions as a function of the transverse momentum of the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and the 4<sup>th</sup> hardest  $b$ -jet at LO and NLO for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  process at the LHC with  $\sqrt{s} = 13$  TeV. The heavy-flavour jets are ordered in  $p_T$ . The upper plots show absolute LO and NLO QCD predictions together with corresponding uncertainty bands. The lower panels display the differential  $\mathcal{K}$ -factor together with the uncertainty band and the relative scale uncertainties of the LO cross section. Results are provided for  $\mu_R = \mu_F = \mu_0 = H_T/3$ . The LO and the NLO NNPDF3.1 PDF sets are employed.

tions are plagued by the same large higher order QCD effects as the integrated fiducial cross sections. Indeed, for example for  $p_T(b_1)$  we observe 90% – 135% NLO QCD corrections, that introduce shape distortions of the order of 45%. The theoretical uncertainties due to scale dependence are, on the other hand, in the range of 20% – 30%. For  $p_T(b_2)$ ,  $p_T(b_3)$  and  $p_T(b_4)$  higher order QCD corrections are of the order of 70% – 120% while uncertainties are between 10% and 25% depending on the  $p_T$  of the  $b$ -jet. Had we used  $\mu_0 = m_t$  instead we would rather obtain shape distortions up to even 150%. This shows the advantage of the dynamical scale over the fixed one.

In figure 6 we show the differential cross section distribution as a function of the angular separation between the  $b$ -jets,  $\Delta R(bb)$ . Also presented are the transverse momentum and invariant mass of the  $bb$  system. They are labelled as  $p_T(bb)$  and  $M(bb)$  respectively. Specifically, we display the two hardest  $b$ -jets, denoted as  $b_1 b_2$ , and the two softest  $b$ -jets,



**Figure 6.** As in figure 5 but for the  $\Delta R(b_1b_2)$ ,  $\Delta R(b_3b_4)$ ,  $p_T(b_1b_2)$ ,  $p_T(b_3b_4)$ ,  $M(b_1b_2)$  and  $M(b_3b_4)$  distributions.

denoted as  $b_3b_4$ . Looking at  $\Delta R(b_1b_2)$  we can notice that the  $b_1b_2$  system originates predominately from top-quark pair production as  $b_1$  and  $b_2$  are generated mostly in back-to-back configurations. This is additionally confirmed by the  $p_T(b_1b_2)$  and  $M(b_1b_2)$  distributions, that have harder spectra in comparison to the  $b_3b_4$  system. The latter system is expected to receive large contributions from gluon splittings as manifested by the enhancement at the beginning of the  $\Delta R(b_3b_4)$  distribution. However, we can also notice rather large contributions in the back-to-back configurations for  $\Delta R(b_3b_4)$ . This suggests that the simple picture that the two high  $p_T$   $b$ -jets are from top-quark decays while the two low  $p_T$   $b$ -jets, which are closest in  $\Delta R(bb)$ , are  $b$ -jets from the  $g \rightarrow b\bar{b}$  splitting may not apply. The reconstruction of the production mechanisms for all final states is rather cumbersome when multiple  $b$ -jets are present. It requires good reconstruction techniques and excellent understanding of the modelling of top-quark decays. The presence of the additional contributions either from off-shell top quarks or from additional resolved light and/or  $b$ -jets makes this picture even more complicated. As we have mentioned earlier we leave such studies on the identification of the origin of the  $b$ -jets for the future. Instead, in the following we will focus on the size of NLO QCD corrections to various observables constructed for the  $b_1b_2$  and  $b_3b_4$  system. We underline here the fact that such spectra are measured experimentally, see e.g. ref. [23]. When examining dimensionful observables we notice that the  $b_1b_2$  system receives larger corrections up to even 150%. For  $b_3b_4$ , on the other hand, corrections are more constant and between 80% and 110%. For dimensionless angular distributions NLO QCD effects are within the 80% – 100% range. All observables, except for  $p_T(b_1b_2)$ , have moderate shape distortions. For the  $p_T(b_1b_2)$  distribution, however, they are even up to 80%. Had we used the fixed scale choice also in this case we would rather obtain more than 100% changes in the shape of the observables due to QCD higher order corrections. Furthermore, for all observables moderate theoretical uncertainties up to 25% – 30% are observed.

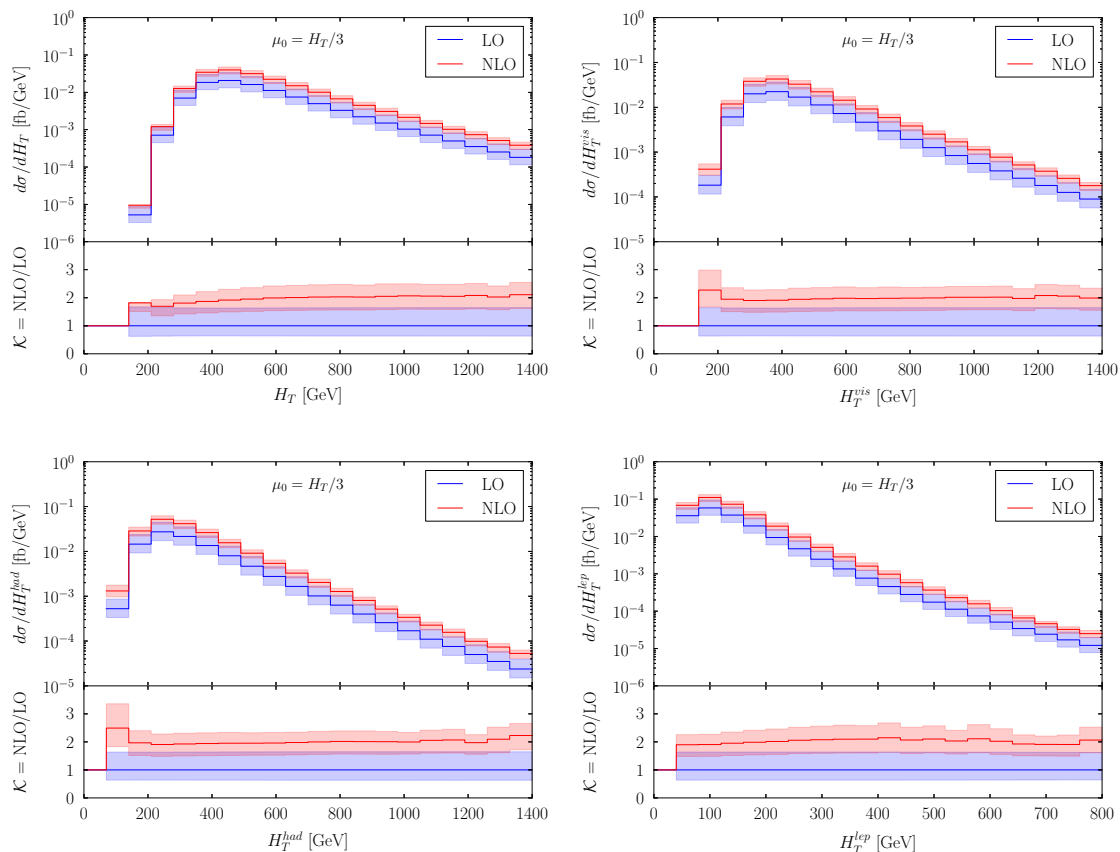
An interesting set of observables is depicted in figure 7. They are the scalar sums built out of the transverse momenta of the various final states from the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process. In particular, we present  $H_T$ ,  $H_T^{\text{vis}}$ ,  $H_T^{\text{had}}$  and  $H_T^{\text{lep}}$ . The first one has already been defined as

$$H_T = \sum_{i=1}^4 p_T(b_i) + \sum_{i=1}^2 p_T(\ell_i) + p_T^{\text{miss}}, \quad (5.1)$$

where  $\ell_{1,2} = e^+, \mu^-$  and  $p_T^{\text{miss}}$  is the missing transverse momentum from the two neutrinos. We also have the visible, hadronic and leptonic versions of  $H_T$ . The three observables are measured experimentally and defined as follows

$$H_T^{\text{vis}} = \sum_{i=1}^4 p_T(b_i) + \sum_{i=1}^2 p_T(\ell_i), \quad H_T^{\text{had}} = \sum_{i=1}^4 p_T(b_i), \quad H_T^{\text{lep}} = \sum_{i=1}^2 p_T(\ell_i). \quad (5.2)$$

These observables have various kinematical thresholds. For example for  $H_T$  and  $H_T^{\text{vis}}$  we have  $H_{T,\text{min}} = H_{T,\text{min}}^{\text{vis}} = 140$  GeV as there is no restriction on the kinematics of the missing transverse momentum. For the remaining two we can write  $H_{T,\text{min}}^{\text{had}} = 100$  GeV and  $H_{T,\text{min}}^{\text{lep}} = 40$  GeV.  $H_T$  observables receive rather constant NLO QCD corrections that are of the order of 100%. Had we changed  $\mu_0 = H_T/3$  to the fixed scale setting

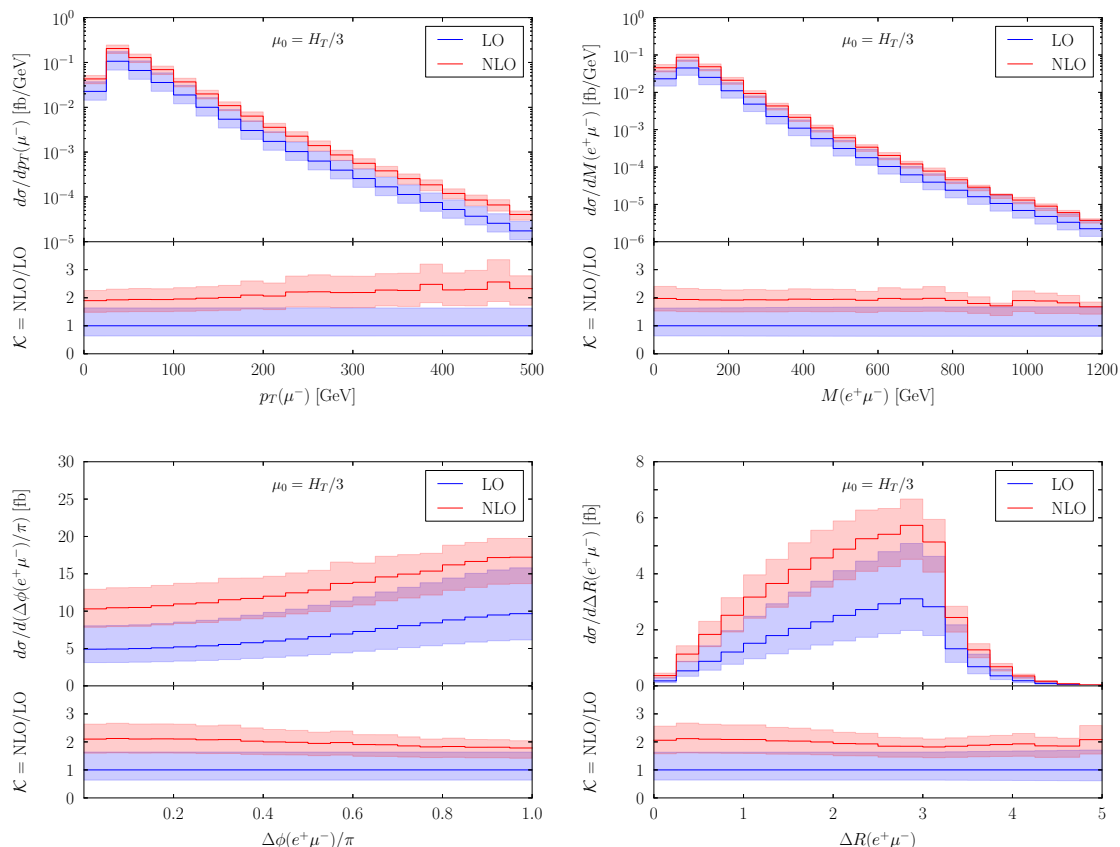


**Figure 7.** As in figure 5 but for the  $H_T$ ,  $H_T^{\text{vis}}$ ,  $H_T^{\text{had}}$  and  $H_T^{\text{lep}}$  distributions.

shape distortions up to almost 200% would rather be noticed. For all four versions of  $H_T$  theoretical uncertainties are of the order of 20% – 30%.

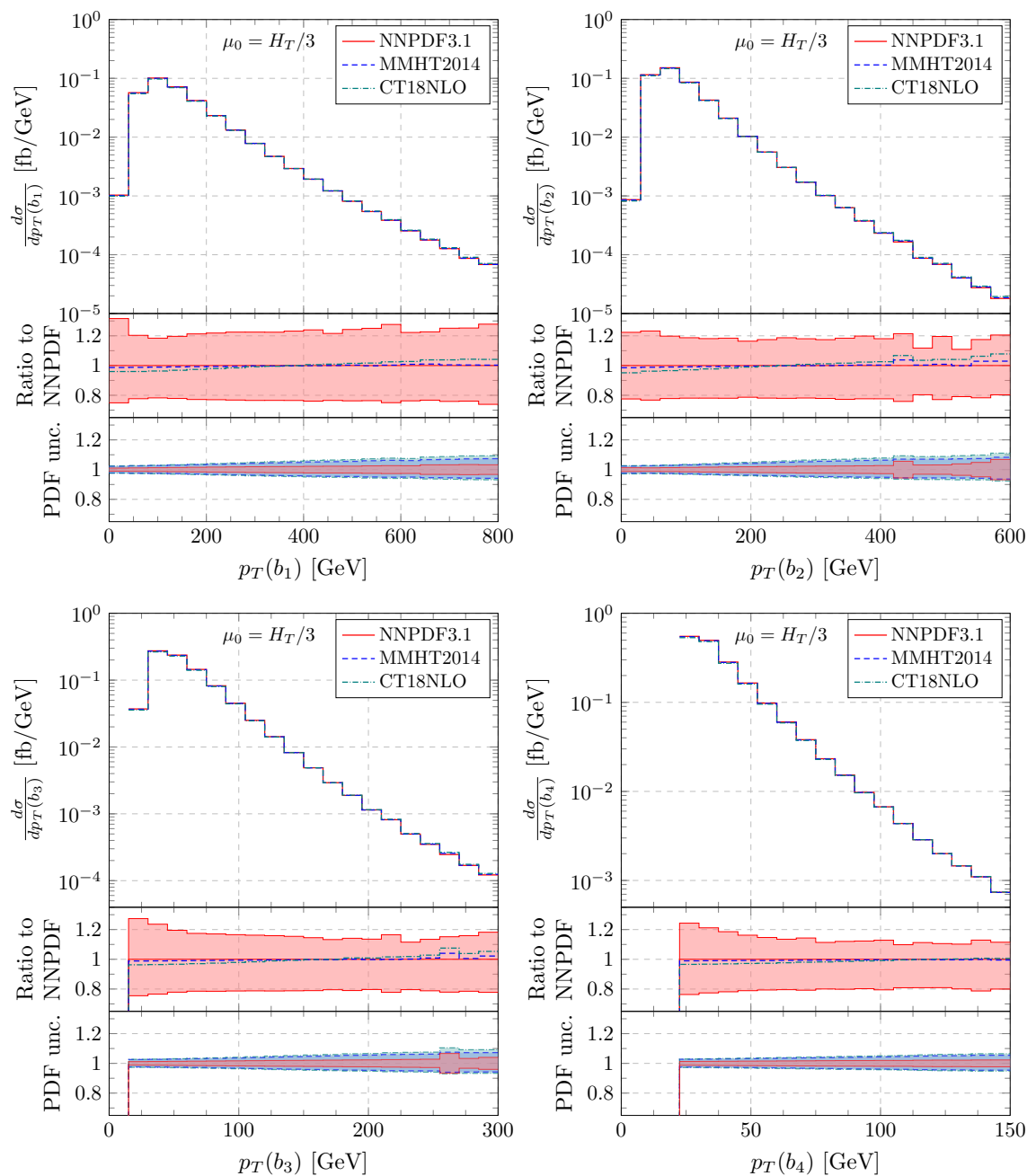
Although the main emphasis is on the understanding of the  $b$ -jet kinematics, we can calculate higher order QCD corrections to any IR-safe observable, which can be constructed from the available final states, in the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process. Therefore, we can examine in detail for example the two charged leptons. The advantage of the leptonic observables over hadronic ones lies in the fact that measurements of lepton observables are particularly precise at the LHC due to the excellent lepton energy resolution of the ATLAS and CMS detectors. In figure 8 we present the following leptonic observables: the transverse momentum of the muon ( $p_T(\mu^-)$ ), the two leptons’ invariant mass ( $M(e^+\mu^-)$ ), angular difference in the transverse plane ( $\Delta\phi(e^+\mu^-)$ ), and angular separation ( $\Delta R(e^+\mu^-)$ ). NLO QCD corrections to leptonic observables are also substantial in the range of 80% – 150%. Theoretical uncertainties, on the other hand, are again up to 20% – 25% only.

As already pointed out the theoretical uncertainty related to the lack of next-to-next-to-leading-order QCD corrections is only a fraction of the overall theoretical systematics. To complete our analysis we study differential cross section uncertainties due to the PDF parameterisation including the latest fits by several groups, i.e. NNPDF3.1, MMHT2014 and CT18NLO. We have comparatively examined the impact of PDFs and scale variations



**Figure 8.** As in figure 5 but for the  $p_T(\mu^-)$ ,  $M(e^+\mu^-)$ ,  $\Delta\phi(e^+\mu^-)$  and  $\Delta R(e^+\mu^-)$  distributions.

on the overall theoretical uncertainty for differential cross section distributions for the  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b} + X$  process. For all observables that we have examined we could confirm that the PDF uncertainties are small and well below the theoretical uncertainties predicted by scale variation. As an example in figure 9 we present NLO differential cross section distributions as a function of the transverse momentum of the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> hardest  $b$ -jet. Upper panels show the absolute NLO predictions for three different PDF sets as obtained with the dynamical scale setting. Middle panels display the relative scale uncertainties of NLO predictions for the NNPDF3.1 PDF set. Also presented are the relative predictions for the MMHT2014 and CT18NLO PDF set. They are obtained by normalisation to the central NLO prediction with the NNPDF3.1 PDF set. Finally, lower panels display the relative internal PDF uncertainties of the NLO cross section for each of the three PDF sets. Also here all NLO predictions are normalised to the central NLO prediction with the NNPDF3.1 PDF set. A few comments are in order. To start with, at the differential level the internal PDF uncertainties are up to 11% for CT18NLO and of the order of 8% for MMHT2014. For the default NNPDF3.1 PDF set they are in the 1%–7% range. When comparing the NLO QCD results as obtained with either CT18NLO or MMHT2014 to the findings generated for the default NNPDF3.1 PDF set the relative differences are up to 8% and 3% respectively. Thus, also at the differential level the internal



**Figure 9.** NLO differential cross section distributions as a function of  $p_T(b_1)$ ,  $p_T(b_2)$ ,  $p_T(b_3)$  and  $p_T(b_4)$ . Results are shown for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  process at the LHC with  $\sqrt{s} = 13$  TeV. Three different PDF sets are employed. Lower panels display the theoretical uncertainties from scales and PDFs. All predictions are normalised to the central NLO prediction with the NNPDF3.1 PDF set.

NNPDF3.1 PDF uncertainties are similar in size to differences between the PDF sets that are recommended for applications at the LHC Run II [64].

To summarise this part, for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  production process at the LHC for a centre-of-mass-system energy of  $\sqrt{s} = 13$  TeV and with rather inclusive selection cuts not only big NLO QCD corrections but also significant shape changes are visible when going from LO to NLO. This confirms that NLO QCD effects to this process are extremely important. The theoretical uncertainties due to scale dependence for  $\mu_0 = H_T/3$  are rather moderate of the order of 20% – 30%. For the fixed scale setting they are much higher. The uncertainties due to the NNPDF3.1 PDF parameterisation are small, i.e. in the 1% – 7% range. Overall, when various PDF sets are examined the PDF uncertainties are maximally up to 11%. Consequently, the dominant component of the final theoretical error for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process is determined by the scale dependence.

## 6 Contribution of initial state bottom quarks

In the next step we study the contributions that are induced by the initial state bottom quarks. To this end, additional subprocesses are included in the calculation. For the LO part we add  $b\bar{b} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b}$ ,  $bb \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b}$  and  $\bar{b}\bar{b} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b}$ . The three reactions are related by crossing symmetry, thus, each comprises 2790 Feynman diagrams. The last two subprocesses must be taken into account as they might not necessarily be suppressed other than by PDFs. Indeed, they are already part of the double-resonant contribution to the  $t\bar{t}$  process with additional two  $b$ -jets. For the real emission part, on the other hand, the following subprocesses are included  $b\bar{b} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} g$ ,  $gb \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} b$ ,  $g\bar{b} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} \bar{b}$ ,  $bb \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} g$  and  $\bar{b}\bar{b} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} g$ . Each subprocess comprises 28728 Feynman diagrams. We continue to use the *anti* -  $k_T$  jet algorithm. However a small modification is needed for the jet flavour assignment. As we are dealing with massless bottom quarks from the theoretical point of view the important parton recombination rules, that are required to guarantee the IR-safety of the jet algorithm, are  $bg \rightarrow b$ ,  $\bar{b}g \rightarrow \bar{b}$  and  $b\bar{b} \rightarrow g$ . We need, however, additional recombination rules for  $bb$  and  $\bar{b}\bar{b}$ . We employ two variants that are IR-safe at NLO. Beyond NLO IR-safety requires the algorithms of refs. [65, 66].

**Charge-blind  $b$ -jet tagging.** In the first case we use the charge-blind  $b$ -jet tagging. From the experimental point of view  $b$ -jet tagging algorithms are sensitive mostly to the absolute flavour and they do not additionally tag the charge of the  $b$ -jet. In the absence of charge tagging any combination that contains an even number of  $b$  and/or  $\bar{b}$  quarks should also be considered to carry zero flavour as from the experimental point of view such signatures will not be distinguishable from  $b\bar{b} \rightarrow g$ . In this case the complete set of recombination rules for heavy-flavour jets is given by

$$bg \rightarrow b, \quad \bar{b}g \rightarrow \bar{b}, \quad b\bar{b} \rightarrow g, \quad bb \rightarrow g, \quad \bar{b}\bar{b} \rightarrow g. \quad (6.1)$$

We ask for at least four  $b$ -jets in the final state and check whether there is at least one combination of (any) four  $b$ -jets that fulfils the required cuts. We shall refer to this approach as the charge-blind  $b$ -jet tagging or in short as the *charge-blind scheme* in the following.



| SCALE           | $\sigma_{gg}^{\text{LO}}$ [fb] | $\sigma_{q\bar{q}+\bar{q}q}^{\text{LO}}$ [fb] | $\sigma_{b\bar{b}+\bar{b}b}^{\text{LO}}$ [fb] | $\sigma_{bb+\bar{b}\bar{b}}^{\text{LO}}$ [fb] |
|-----------------|--------------------------------|---|---|---|
| $\mu_0 = m_t$   | 6.561(2)                       | 0.4367(1)                                     | 0.008607(7)                                   | 0.006184(8)                                   |
| $\mu_0 = H_T/3$ | 6.404(3)                       | 0.4092(1)                                     | 0.008428(7)                                   | 0.006005(7)                                   |

**Table 5.** LO integrated fiducial cross section for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b}b\bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV for  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$ . Theoretical results with and without the initial state  $b$  contributions are provided for the LO NNPDF3.1 PDF set. Also given are Monte Carlo integration errors (in parenthesis).

**Charge-aware  $b$ -jet tagging.** In the second case, we assume that the charge tagging of  $b$ -jets is possible, see e.g. refs. [67–70]. From the experimental point of view the disadvantage of this approach might lie in the possibility of the reduction in the  $b$ -jet tagging efficiency and in smaller event statistics. In this case the following recombinations rules are employed:

$$bg \rightarrow b, \quad \bar{b}g \rightarrow \bar{b}, \quad b\bar{b} \rightarrow g, \quad bb \rightarrow b, \quad \bar{b}\bar{b} \rightarrow \bar{b}. \quad (6.2)$$

We ask for at least four  $b$ -jets in the final state, however, this time we check whether any combination with zero total charge, i.e. any  $b\bar{b}b\bar{b}$  combination, passes the cuts. In this scheme there is no need to consider the  $bb$  and  $\bar{b}\bar{b}$  initiated subprocess. We shall refer to this approach as the charge-aware  $b$ -jet tagging or in short as the *charge-aware scheme*.

In the following the difference between the two approaches is examined. We investigate their impact on the full  $pp$  cross sections at LO and NLO in QCD. Specifically, we compare the two schemes to the case when the initial state bottom-quark contribution is omitted. At LO the two approaches are equivalent because we always have exactly the four  $b$ -jets that need to pass the cuts. Furthermore, at this order both schemes differ only by the initial state subprocesses that must be taken into account. Simply, in the charge-aware scheme the  $bb$  and  $\bar{b}\bar{b}$  subprocesses are not considered.

In table 5 we present contributions to the LO  $t\bar{t}b\bar{b}$  cross section subprocess by subprocess. As already stipulated  $\sigma_{b\bar{b}+\bar{b}b}^{\text{LO}}$  and  $\sigma_{bb+\bar{b}\bar{b}}^{\text{LO}}$  are similar in size. For  $\mu_0 = m_t$  the full  $pp$  LO integrated cross section with the  $b$ -initiated contributions included is given by

$$\sigma_{\text{charge-aware}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 7.006(2) \text{ fb}, \quad (6.3)$$

for the charge-aware scheme and reads

$$\sigma_{\text{charge-blind}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 7.012(2) \text{ fb}, \quad (6.4)$$

for the charge-blind one. The two findings can be compared to the result where the initial state  $b$  contributions are neglected

$$\sigma_{\text{no-}b}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = m_t) = 6.998(2) \text{ fb}. \quad (6.5)$$

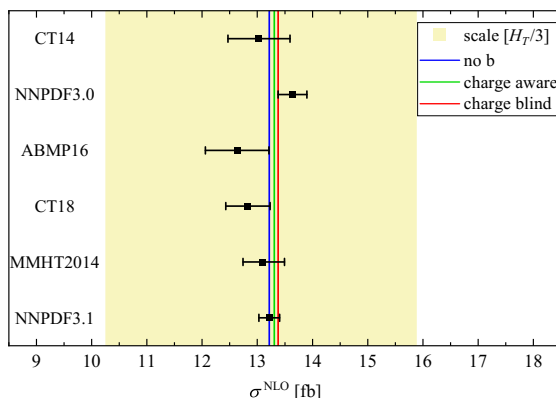
Similar results are obtained for  $\mu_0 = H_T/3$ . We can conclude that at LO the initial state  $b$ -quark contributions are at the 0.1% – 0.2% level independently of the scale choice and the  $b$ -jet tagging scheme. Thus, they can be safely neglected.

| $p_T^{\text{veto}}(j)$          | $p_T(b)$ | $\sigma_{\text{no } b}^{\text{NLO}}$ [fb] | $\sigma_{\text{ch-aware}}^{\text{NLO}}$ [fb] | $\sigma_{\text{ch-blind}}^{\text{NLO}}$ [fb] | $\delta_{\text{scale}}$ | $\delta_{\text{PDF}}$ |
|---------------------------------|----------|---|--|--|-------------------------|-----------------------|
| $\mu_R = \mu_F = \mu_0 = m_t$   |          |   |  |  |                         |                       |
|                                 | 25       | 13.24(3)                                  | 13.33(3)                                     | 13.41(3)                                     | +18%<br>-22%            | +1%<br>-1%            |
|                                 | 30       | 9.25(2)                                   | 9.32(2)                                      | 9.37(2)                                      | +14%<br>-21%            | +2%<br>+2%            |
|                                 | 35       | 6.57(1)                                   | 6.62(1)                                      | 6.66(1)                                      | +12%<br>-20%            | +2%<br>-2%            |
|                                 | 40       | 4.70(1)                                   | 4.74(1)                                      | 4.77(1)                                      | +10%<br>-19%            | +2%<br>-2%            |
| 100                             | 25       | 10.37(3)                                  | 10.46(3)                                     | 10.53(3)                                     | +3%<br>-16%             | +1%<br>-1%            |
| 50                              | 25       | 7.77(3)                                   | 7.85(3)                                      | 7.93(3)                                      | +3%<br>-33%             | +1%<br>-1%            |
| $\mu_R = \mu_F = \mu_0 = H_T/3$ |          |   |  |  |                         |                       |
|                                 | 25       | 13.22(3)                                  | 13.31(3)                                     | 13.38(3)                                     | +20%<br>-22%            | +1%<br>-1%            |
|                                 | 30       | 9.09(2)                                   | 9.16(2)                                      | 9.21(2)                                      | +18%<br>-22%            | +2%<br>-2%            |
|                                 | 35       | 6.37(1)                                   | 6.42(1)                                      | 6.46(1)                                      | +17%<br>-21%            | +2%<br>-2%            |
|                                 | 40       | 4.51(1)                                   | 4.54(1)                                      | 4.57(1)                                      | +16%<br>-21%            | +2%<br>-2%            |
| 100                             | 25       | 10.77(3)                                  | 10.86(3)                                     | 10.94(3)                                     | +8%<br>-18%             | +1%<br>-1%            |
| 50                              | 25       | 8.35(3)                                   | 8.43(3)                                      | 8.51(3)                                      | +1%<br>-18%             | +1%<br>-1%            |

**Table 6.** NLO integrated fiducial cross section for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV for  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$ . Theoretical results with and without the initial state  $b$  contributions are provided for the NLO NNPDF3.1 PDF set. Results for four different values of the  $p_T(b)$  cut and for a jet veto with  $p_T^{\text{veto}}(j) = 50, 100$  GeV are additionally provided. Also given are the theoretical uncertainties due to scale variation ( $\delta_{\text{scale}}$ ) and PDFs ( $\delta_{\text{PDF}}$ ). Finally, Monte Carlo integration errors are shown (in parenthesis).

In table 6 we display our findings at NLO in QCD. The theoretical results are given for both scale choices  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$ . For  $\mu_0 = m_t$  the initial state bottom-quark contributions are of the order of 1.3%. The variation in the  $p_T(b)$  cut in the range of 25–40 GeV has increased these contributions up to 1.5%. A similar increase can be obtained with the imposition of a jet veto on the fifth jet (upper bound on the allowed transverse momentum) of 100 GeV. On the other hand, for  $p_T^{\text{veto}}(j) = 50$  GeV these contributions have reached 2%. Results for the dynamical scale choice are very similar. Consequently, the bottom-quark initiated contributions can be safely omitted when compared to the theoretical error for this process. The latter is at the 20% – 22% level.

A few comments are in order. If we look at the bottom-quark initiated subprocesses only, their combined contribution at NLO is dominated by the  $bg/\bar{b}g$  initial states. Furthermore, at this level there are substantial differences between the two schemes. Specifically,



**Figure 10.** NLO integrated fiducial cross section for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV for  $\mu_0 = H_T/3$ . Theoretical results with and without the initial state  $b$ -quark contributions are provided for the NLO NNPDF3.1 PDF set. Also given are results with other PDF sets for the case with no initial state bottom-quark contributions.

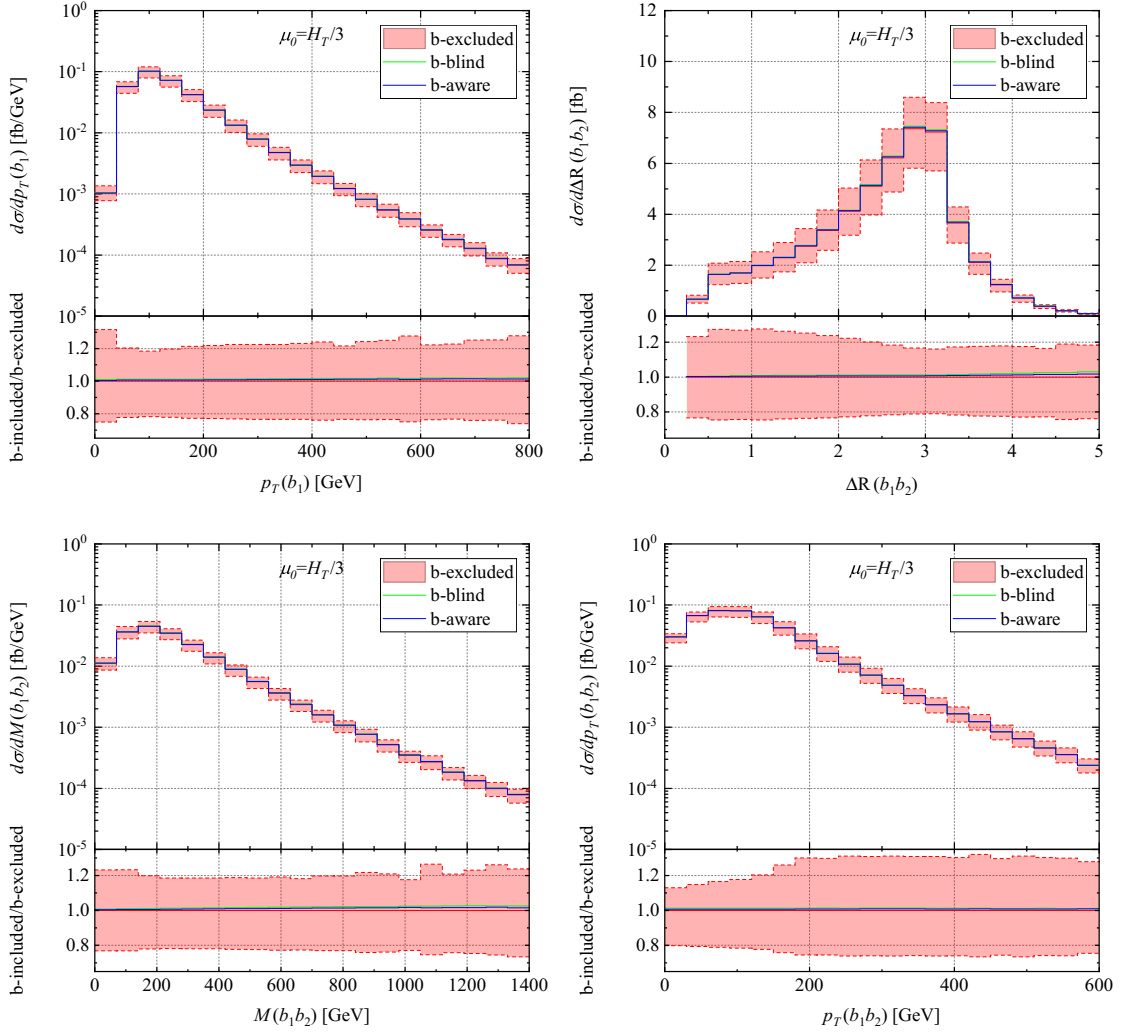
for  $\mu_0 = m_t$  we have for the charge-aware and charge-blind scheme the following results

$$\begin{aligned} \sigma_{b\text{-initiated, charge-aware}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) &= 0.0954(2) \text{ fb} \\ \sigma_{b\text{-initiated, charge-blind}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = m_t) &= 0.1684(4) \text{ fb}, \end{aligned} \tag{6.6}$$

where in bracket the MC error is additionally provided. It is only when the dominant  $gg$  partonic subprocess is taken into account that the differences between the two schemes become insignificant. Furthermore, having at hand the results with an additional jet veto on the fifth jet we can make the following remarks on the size of the higher order QCD effects. We observe that the jet veto can reduce the  $\mathcal{K}$ -factor for the default set of cuts. Specifically, for  $\mu_0 = m_t$  ( $\mu_0 = H_T/3$ ) when  $p_T^{\text{veto}}(j) = 100$  GeV is used we obtain  $\mathcal{K} = 1.48$  ( $\mathcal{K} = 1.58$ ), whereas for  $p_T^{\text{veto}}(j) = 50$  GeV we have  $\mathcal{K} = 1.11$  ( $\mathcal{K} = 1.23$ ).

As a bonus of our study in figure 10 we show graphically NLO integrated fiducial cross section for the  $pp \rightarrow t \bar{t} b \bar{b}$  process with and without the initial state  $b$  contributions. Results are given for the NLO NNPDF3.1 PDF set. However, we also provide results with other PDF sets for the case when the initial state  $b$ -quark contributions are not included. We display the result for CT14, CT18, NNPDF3.0, MMHT2014 and for the NLO ABMP16 PDF set [71]. In the latter case the internal PDF uncertainties are evaluated using the symmhessian method, see e.g. ref. [58]. To put all the results in perspective the theoretical uncertainty resulting from scale dependence are also shown. All predictions are calculated for  $\mu_0 = H_T/3$ . We notice that, all theoretical predictions agree within their internal PDF errors with the default result for the NNPDF3.1 PDF set. The ABMP16 PDF uncertainties are, however, the largest, i.e. at the 5% level. Moreover, the earlier version of CT and NNPDF PDF sets, i.e. CT14 and NNPDF3.0, show larger uncertainties than those given by CT18 and NNPDF3.1. For CT14 we obtained 4% while for NNPDF3.0 2%. This should be compared to 3% and 1% for CT18 and NNPDF3.1 respectively.

Finally, we perform a similar comparative analysis at the differential level. In figure 11 we display NLO differential cross section distributions as a function of  $p_T(b_1)$ ,  $\Delta R(b_1 b_2)$ ,



**Figure 11.** Differential cross section distributions as a function of  $p_T(b_1)$ ,  $\Delta R(b_1 b_2)$ ,  $M(b_1 b_2)$  and  $p_T(b_1 b_2)$  for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b \bar{b} b \bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV. The upper plots show absolute NLO QCD predictions without and with the initial state bottom quark contributions. In the latter case results are shown for the charge blind and charge aware  $b$ -jet tagging. The lower panels display the differential ratios of these contributions. Also shown is the NLO scale dependence for the case without the initial state bottom quark contributions. The NLO NNPDF3.1 PDF set is employed and  $\mu_0 = H_T/3$  is used.

$M(b_1 b_2)$  and  $p_T(b_1 b_2)$ . The upper plots show absolute NLO QCD predictions with and without the initial state  $b$ -quark contributions. Also shown is the NLO scale dependence for the default case. The lower panels display the differential ratios of the results with the initial state bottom-quark contributions to the one without such contributions. Overall, for the process at hand when comparing to the differential theoretical errors due to scale dependence the inclusion of the bottom-quark induced subprocesses is not important also at the differential level. Nonetheless, a consistent treatment of heavy-flavour jets is necessary to obtain IR-finite results.

## 7 Comparison with previous results

In what follows we compare our predictions for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} + X$  process at the LHC running at  $\sqrt{s} = 13$  TeV to the theoretical predictions from ref. [25], which we dub as DLP computation. The comparison is carried out at the integrated and differential level at LO and NLO in QCD. We adopt the dynamical scale setting used in ref. [25]. Specifically, we employ  $\mu_R = \mu_F = \mu_0$  where  $\mu_0 = \mu_{\text{DLP}}$  is given by

$$\mu_0 = \mu_{\text{DLP}} = \frac{1}{2} \left[ \left( p_T^{\text{miss}} + \sum_{i=e^+, \mu^-, b_1, b_2, b_3, b_4} E_T(i) \right) + 2m_t \right]^{1/2} \left( \sum_{i=b_1, b_2, b_3, b_4} E_T(i) \right)^{1/2}. \quad (7.1)$$

Furthermore,  $E_T(i) = \sqrt{p_T^2(i) + M^2(i)}$  and  $M^2(i)$  is the invariant mass squared of the object considered. Similarly to our dynamical scale setting also in this case the top-quark reconstruction is not attempted. In ref. [25] the bottom-quark induced contributions are included at LO and NLO in QCD. The scheme adopted there is that of the charge-blind  $b$ -jet tagging, where the  $bb$  and  $\bar{b}\bar{b}$  subprocesses are neglected. We start with the LO results for the integrated fiducial cross section. When the  $bb$  and  $\bar{b}\bar{b}$  contributions are neglected the LO integrated fiducial cross section as obtained with the help of HELAC-NLO is given by

$$\sigma_{\text{HELAC-NLO}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = \mu_{\text{DLP}}) = 5.201(2)_{-35\%}^{+60\%} \text{ fb}. \quad (7.2)$$

It can be directly compared to the result from ref. [25]

$$\sigma_{\text{DLP}}^{\text{LO}}(\text{NNPDF3.1}, \mu_0 = \mu_{\text{DLP}}) = 5.198(4)_{-35\%}^{+60\%} \text{ fb}. \quad (7.3)$$

The two results agree perfectly within the given MC errors. Had we additionally included the two missing subprocesses  $bb$  and  $\bar{b}\bar{b}$  the result would rather be  $\sigma_{\text{HELAC-NLO}}^{\text{LO}} = 5.206(2) \text{ fb}$ . The latter is  $1.8\sigma$  away from  $\sigma_{\text{DLP}}^{\text{LO}}$ .

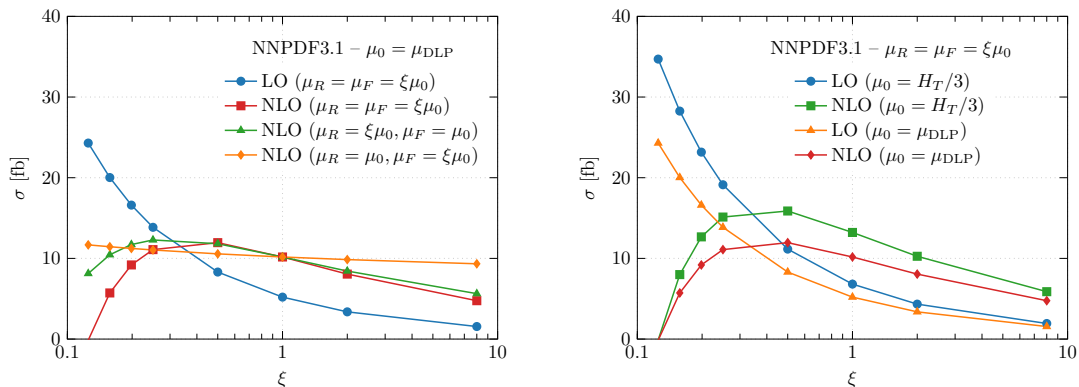
In order to reproduce the NLO QCD result from ref. [25] the additional fifth jet (if resolved) must be added to the definition of  $\mu_0 = \mu_{\text{DLP}}$ . The  $\mu_0 = \mu_{\text{DLP}}$  scale is, thus, given by

$$\mu_0 = \mu_{\text{DLP}} = \frac{1}{2} \left[ \left( p_T^{\text{miss}} + \sum_{i=e^+, \mu^-, b_1, b_2, b_3, b_4, j} E_T(i) \right) + 2m_t \right]^{1/2} \left( \sum_{i=b_1, b_2, b_3, b_4, j} E_T(i) \right)^{1/2}. \quad (7.4)$$

The HELAC-NLO NLO result and the  $\sigma_{\text{DLP}}^{\text{NLO}}$  one, both without the  $bb$  and  $\bar{b}\bar{b}$  contributions, read

$$\begin{aligned} \sigma_{\text{HELAC-NLO}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = \mu_{\text{DLP}}) &= 10.28(1)_{-21\%}^{+18\%} \text{ fb}, \\ \sigma_{\text{DLP}}^{\text{NLO}}(\text{NNPDF3.1}, \mu_0 = \mu_{\text{DLP}}) &= 10.28(8)_{-21\%}^{+18\%} \text{ fb}. \end{aligned} \quad (7.5)$$

Theoretical predictions are in perfect agreement. Had we also included the  $bb$  and  $\bar{b}\bar{b}$  contributions at NLO our result would rather be  $\sigma_{\text{HELAC-NLO}}^{\text{NLO}} = 10.29(1) \text{ fb}$ , so still in agreement with  $\sigma_{\text{DLP}}^{\text{NLO}}$ .



**Figure 12.** Scale dependence of the LO and NLO integrated fiducial cross sections for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV for  $\mu_0 = \mu_{\text{DLP}}$ . The LO and NLO NNPDF3.1 PDF sets are employed. Also shown is the variation of  $\mu_R$  with fixed  $\mu_F$  and the variation of  $\mu_F$  with fixed  $\mu_R$  as well as the comparison to our default scale choice  $\mu_0 = H_T/3$ .

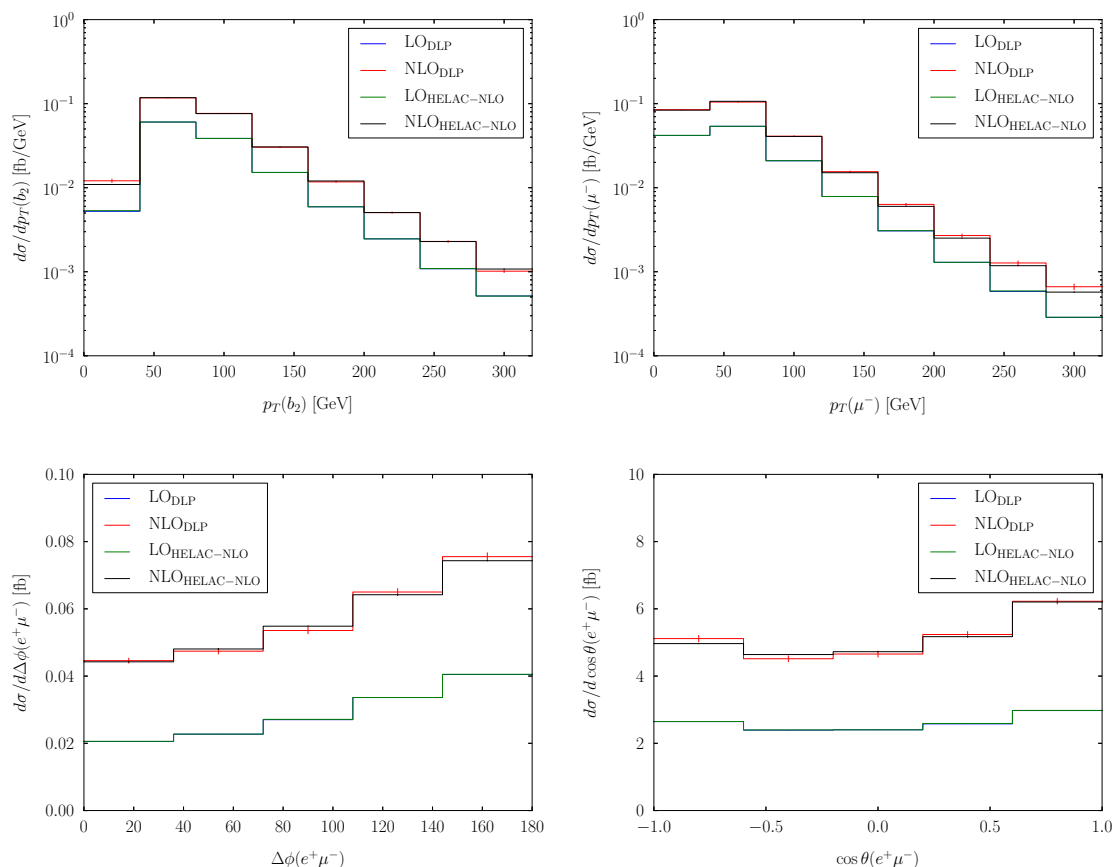
A few comments are in order. Both LO and NLO QCD results for  $\mu_0 = \mu_{\text{DLP}}$  are much smaller than these obtained either for  $\mu_0 = m_t$  or  $\mu_0 = H_T/3$ . In the case of the NLO predictions the differences exceed the theoretical uncertainties due to scale dependence. The dynamical scale  $\mu_0 = \mu_{\text{DLP}}$  includes the dependence on the  $p_T$  of the additional jet, thus, on the average its value is larger. On the other hand, the asymptotic freedom guarantees that the value of  $\alpha_s$  becomes smaller for larger  $\mu_0$ , resulting in lower NLO and LO cross sections. Had we removed the additional jet from the definition of  $\mu_0 = \mu_{\text{DLP}}$  the resulting NLO cross section would increase by about 13%. Consequently, it would be in agreement with the NLO findings both for  $\mu_0 = m_t$  and  $\mu_0 = H_T/3$  within the quoted theoretical uncertainties.

In figure 12 we provide the graphical representation of the scale dependence of the LO and NLO integrated fiducial cross sections for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  production process as obtained with  $\mu_0 = \mu_{\text{DLP}}$ . Also shown are the variation of  $\mu_R$  with fixed  $\mu_F$  and the variation of  $\mu_F$  with fixed  $\mu_R$ . Finally, the comparison to our dynamical scale setting of  $\mu_R = \mu_F = \mu_0 = H_T/3$  is depicted.

In figure 13 we depict the comparison between the HELAC-NLO predictions and the theoretical results from ref. [25] at the differential level. Specifically, we show LO and NLO differential cross section distributions as a function of  $p_T(b_2)$ ,  $p_T(\mu^-)$ ,  $\Delta\phi(e^+ \mu^-)$  and  $\cos\theta(e^+ \mu^-)$ . Also here heavy-flavour induced subprocesses are included except from  $bb$  and  $b\bar{b}$ . For each observable good agreement has been found.

## 8 Summary

In this paper we provided state-of-the-art theoretical predictions for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process at the LHC with  $\sqrt{s} = 13$  TeV. In the computation off-shell top quarks were described by Breit-Wigner distributions. Furthermore double-, single- as well as non-resonant top-quark contributions and interference effects were consistently in-



**Figure 13.** Differential cross section distributions as a function of  $p_T(b_2)$ ,  $p_T(\mu^-)$ ,  $\Delta\phi(e^+\mu^-)$  and  $\cos\theta(e^+\mu^-)$  for the  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}b\bar{b} + X$  production process at the LHC with  $\sqrt{s} = 13$  TeV. Results are presented at LO and NLO for  $\mu_0 = \mu_{\text{DLP}}$ . Comparison between HELAC-NLO predictions and results from ref. [25] (given with the subscript DLP) is shown. The NNPDF3.1 PDF sets are employed. Also displayed are Monte Carlo integration errors.

incorporated at the matrix element level. Non-resonant and off-shell effects due to the finite  $W$ -boson width were also consistently taken into account. All results were obtained with the help of the HELAC-NLO MC package. We have shown LO and NLO predictions for the integrated and differential cross sections, that are phenomenologically relevant for LHC physics. We assessed the theoretical uncertainties of our high-precision theoretical predictions stemming from scale dependence and PDF parameterisation while using fixed and dynamical scale settings, i.e.  $\mu_R = \mu_F = \mu_0$ , where  $\mu_0 = m_t$  or  $\mu_0 = H_T/3$ . Furthermore, the following PDF sets were examined NNPDF3.1, CT18 and MMHT14.

The full  $pp$  cross section receives positive and large NLO QCD corrections of 89%. The theoretical uncertainties resulting from scale variation are 65% at LO and 22% at NLO. The internal PDF uncertainties are very small, at the level of 1% – 3% only. Thus, the NLO theoretical error is completely dominated by the scale dependence. We provide LO and NLO results for other PDF sets and observe that the  $\mathcal{K}$ -factor has a very large spread from 1.81 down to 1.23 depending on the LO PDF set employed and specifically on the

value of the strong coupling constant  $\alpha_s(m_Z)$  used. On the other hand, we observed a very stable behaviour of the systematics when varying the  $p_T(b)$  cut or adding additional cuts.

The differential cross sections have been plagued by the same large higher order QCD effects as the integrated fiducial cross sections. Not only big NLO QCD corrections but also significant shape changes were visible when going from LO to NLO. This confirms that NLO QCD effects to the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process are extremely important. The theoretical uncertainties due to scale dependence for  $\mu_0 = H_T/3$  are rather moderate of the order of 20% – 30%. For the fixed scale setting they are much higher. The uncertainties due to the NNPDF3.1 PDF parameterisation are small, i.e. in the 1% – 7% range. When other PDF sets have been examined the PDF uncertainties increased maximally up to 11%. Consequently, the final theoretical error for the process at hand remains dominated by the scale dependence.

In the next step we have studied the contributions that are induced by the initial state bottom quarks. To this end, additional subprocesses were included in the calculation. Additionally, we needed new recombination rules for partons to construct light- and heavy-flavour jets. We employed two variants that are IR-safe at NLO: charge-blind and charge-aware  $b$ -jet tagging. The differences between the two approaches were examined in detail. Overall, for the process at hand when comparing to the theoretical errors due to scale dependence the size of the contributions generated by the bottom-quark induced subprocesses is negligible both at the integrated and differential level. Nonetheless, a consistent treatment of heavy-flavour jets is necessary to obtain IR-finite results.

Finally, we compared our predictions for the  $pp \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b} b\bar{b} + X$  process to the previous results obtained in ref. [25]. After clarifying the scale choice used in ref. [25] with the authors and the status for the two  $b\bar{b}$  and  $b\bar{b}$  subprocesses perfect agreement has been found both at the integrated and differential level.

## Acknowledgments

We would like to thank Jasmina Nasufi for her contributions at early stages of this project. We would like to thank Mathieu Pellen for clarifications concerning the dynamical scale setting and the setup used in ref. [25]. Furthermore, we would like to thank him for providing the data files for differential cross sections distributions that we have used for comparisons in section 7.

The work of H.Y.B. and M.W. was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant 396021762 – TRR 257: *P3H – Particle Physics Phenomenology after the Higgs Discovery*. Support by a grant of the Bundesministerium für Bildung und Forschung (BMBF) is additionally acknowledged.

The research of M.L. was supported by the DFG under grant 400140256 – GRK 2497: *The physics of the heaviest particles at the Large Hardon Collider*.

The work of G.B. was supported by grant K 125105 of the National Research, Development and Innovation Office in Hungary.

H.B.H. has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (grant agreement



no. 683211). Furthermore, the work of H.B.H has been partially supported by STFC consolidated HEP theory grant ST/T000694/1

The work of M.K. was supported in part by the U.S. Department of Energy under grant DE-SC0010102.

Simulations were performed with computing resources granted by RWTH Aachen University under projects `rwth0414` and `rwth0531`.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] CMS collaboration, *Observation of  $t\bar{t}H$  production*, *Phys. Rev. Lett.* **120** (2018) 231801 [[arXiv:1804.02610](https://arxiv.org/abs/1804.02610)] [[INSPIRE](#)].
- [2] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [[arXiv:1806.00425](https://arxiv.org/abs/1806.00425)] [[INSPIRE](#)].
- [3] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922](https://arxiv.org/abs/1610.07922) [[INSPIRE](#)].
- [4] CMS collaboration, *Search for  $t\bar{t}H$  production in the  $H \rightarrow b\bar{b}$  decay channel with leptonic  $t\bar{t}$  decays in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **03** (2019) 026 [[arXiv:1804.03682](https://arxiv.org/abs/1804.03682)] [[INSPIRE](#)].
- [5] ATLAS collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a  $b\bar{b}$  pair in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 072016 [[arXiv:1712.08895](https://arxiv.org/abs/1712.08895)] [[INSPIRE](#)].
- [6] R.V. Harlander, S.Y. Klein and M. Lipp, *FeynGame*, *Comput. Phys. Commun.* **256** (2020) 107465 [[arXiv:2003.00896](https://arxiv.org/abs/2003.00896)] [[INSPIRE](#)].
- [7] T. Plehn, G.P. Salam and M. Spannowsky, *Fat Jets for a Light Higgs*, *Phys. Rev. Lett.* **104** (2010) 111801 [[arXiv:0910.5472](https://arxiv.org/abs/0910.5472)] [[INSPIRE](#)].
- [8] ATLAS collaboration, *Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. D* **99** (2019) 052009 [[arXiv:1811.02305](https://arxiv.org/abs/1811.02305)] [[INSPIRE](#)].
- [9] CMS collaboration, *Search for standard model production of four top quarks in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Lett. B* **772** (2017) 336 [[arXiv:1702.06164](https://arxiv.org/abs/1702.06164)] [[INSPIRE](#)].
- [10] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, *NLO QCD corrections to  $t\bar{t}b\bar{b}$  production at the LHC: 1. Quark-antiquark annihilation*, *JHEP* **08** (2008) 108 [[arXiv:0807.1248](https://arxiv.org/abs/0807.1248)] [[INSPIRE](#)].
- [11] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, *NLO QCD corrections to  $pp \rightarrow t\bar{t}b\bar{b} + X$  at the LHC*, *Phys. Rev. Lett.* **103** (2009) 012002 [[arXiv:0905.0110](https://arxiv.org/abs/0905.0110)] [[INSPIRE](#)].
- [12] G. Bevilacqua, M. Czakon, C.G. Papadopoulos, R. Pittau and M. Worek, *Assault on the NLO Wishlist:  $pp \rightarrow t\bar{t}b\bar{b}$* , *JHEP* **09** (2009) 109 [[arXiv:0907.4723](https://arxiv.org/abs/0907.4723)] [[INSPIRE](#)].

- [13] 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](#)] [[INSPIRE](#)].
- [14] 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](#)].
- [15] 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](#)].
- [16] 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](#)] [[INSPIRE](#)].
- [17] A. Kardos and Z. Trócsányi, *Hadroproduction of  $t\bar{t}$  pair with a  $b\bar{b}$  pair using PowHel*, *J. Phys. G* **41** (2014) 075005 [[arXiv:1303.6291](#)] [[INSPIRE](#)].
- [18] 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](#)] [[INSPIRE](#)].
- [19] 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*, *JHEP* **03** (2015) 083 [[arXiv:1408.0266](#)] [[INSPIRE](#)].
- [20] G. Bevilacqua, M.V. Garzelli and A. Kardos,  *$t\bar{t}b\bar{b}$  hadroproduction with massive bottom quarks with PowHel*, [arXiv:1709.06915](#) [[INSPIRE](#)].
- [21] 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](#)] [[INSPIRE](#)].
- [22] CMS collaboration, *Measurements of  $t\bar{t}$  cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Lett. B* **776** (2018) 355 [[arXiv:1705.10141](#)] [[INSPIRE](#)].
- [23] ATLAS collaboration, *Measurements of inclusive and differential fiducial 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*, *JHEP* **04** (2019) 046 [[arXiv:1811.12113](#)] [[INSPIRE](#)].
- [24] CMS collaboration, *Measurement of the cross section for  $t\bar{t}$  production with additional jets and b jets in pp collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **07** (2020) 125 [[arXiv:2003.06467](#)] [[INSPIRE](#)].
- [25] A. Denner, J.-N. Lang and M. Pellen, *Full NLO QCD corrections to off-shell  $t\bar{t}b\bar{b}$  production*, [arXiv:2008.00918](#) [[INSPIRE](#)].
- [26] G. Bevilacqua et al., *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986 [[arXiv:1110.1499](#)] [[INSPIRE](#)].
- [27] G. Bevilacqua, H.B. Hartanto, M. Kraus and M. Worek, *Top Quark Pair Production in Association with a Jet with Next-to-Leading-Order QCD Off-Shell Effects at the Large Hadron Collider*, *Phys. Rev. Lett.* **116** (2016) 052003 [[arXiv:1509.09242](#)] [[INSPIRE](#)].
- [28] G. Bevilacqua, H.B. Hartanto, M. Kraus and M. Worek, *Off-shell Top Quarks with One Jet at the LHC: A comprehensive analysis at NLO QCD*, *JHEP* **11** (2016) 098 [[arXiv:1609.01659](#)] [[INSPIRE](#)].

- [29] G. Bevilacqua, H.B. Hartanto, M. Kraus, T. Weber and M. Worek, *Hard Photons in Hadroproduction of Top Quarks with Realistic Final States*, *JHEP* **10** (2018) 158 [[arXiv:1803.09916](#)] [[INSPIRE](#)].
- [30] G. Bevilacqua, H.B. Hartanto, M. Kraus, T. Weber and M. Worek, *Towards constraining Dark Matter at the LHC: Higher order QCD predictions for  $t\bar{t} + Z(Z \rightarrow \nu_\ell \bar{\nu}_\ell)$* , *JHEP* **11** (2019) 001 [[arXiv:1907.09359](#)] [[INSPIRE](#)].
- [31] G. Bevilacqua, H.-Y. Bi, H.B. Hartanto, M. Kraus and M. Worek, *The simplest of them all:  $t\bar{t}W^\pm$  at NLO accuracy in QCD*, *JHEP* **08** (2020) 043 [[arXiv:2005.09427](#)] [[INSPIRE](#)].
- [32] M. Czakon, C.G. Papadopoulos and M. Worek, *Polarizing the Dipoles*, *JHEP* **08** (2009) 085 [[arXiv:0905.0883](#)] [[INSPIRE](#)].
- [33] A. van Hameren, *PARNI for importance sampling and density estimation*, *Acta Phys. Polon. B* **40** (2009) 259 [[arXiv:0710.2448](#)] [[INSPIRE](#)].
- [34] A. van Hameren, *Kaleu: A General-Purpose Parton-Level Phase Space Generator*, [arXiv:1003.4953](#) [[INSPIRE](#)].
- [35] A. Denner, S. Dittmaier, M. Roth and D. Wackerth, *Predictions for all processes  $e^+e^- \rightarrow 4$  fermions +  $\gamma$* , *Nucl. Phys. B* **560** (1999) 33 [[hep-ph/9904472](#)] [[INSPIRE](#)].
- [36] A. Denner, S. Dittmaier, M. Roth and L.H. Wieders, *Electroweak corrections to charged-current  $e^+e^- \rightarrow 4$  fermion processes: Technical details and further results*, *Nucl. Phys. B* **724** (2005) 247 [Erratum *ibid.* **854** (2012) 504] [[hep-ph/0505042](#)] [[INSPIRE](#)].
- [37] G. Bevilacqua, M. Czakon, A. van Hameren, C.G. Papadopoulos and M. Worek, *Complete off-shell effects in top quark pair hadroproduction with leptonic decay at next-to-leading order*, *JHEP* **02** (2011) 083 [[arXiv:1012.4230](#)] [[INSPIRE](#)].
- [38] A. Denner, S. Dittmaier, S. Kallweit and S. Pozzorini, *NLO QCD corrections to off-shell top-antitop production with leptonic decays at hadron colliders*, *JHEP* **10** (2012) 110 [[arXiv:1207.5018](#)] [[INSPIRE](#)].
- [39] A. van Hameren, C.G. Papadopoulos and R. Pittau, *Automated one-loop calculations: A Proof of concept*, *JHEP* **09** (2009) 106 [[arXiv:0903.4665](#)] [[INSPIRE](#)].
- [40] G. Ossola, C.G. Papadopoulos and R. Pittau, *CutTools: A Program implementing the OPP reduction method to compute one-loop amplitudes*, *JHEP* **03** (2008) 042 [[arXiv:0711.3596](#)] [[INSPIRE](#)].
- [41] G. Ossola, C.G. Papadopoulos and R. Pittau, *Reducing full one-loop amplitudes to scalar integrals at the integrand level*, *Nucl. Phys. B* **763** (2007) 147 [[hep-ph/0609007](#)] [[INSPIRE](#)].
- [42] P. Nogueira, *Automatic Feynman graph generation*, *J. Comput. Phys.* **105** (1993) 279.
- [43] 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](#)].
- [44] A. van Hameren, *OneLoop: For the evaluation of one-loop scalar functions*, *Comput. Phys. Commun.* **182** (2011) 2427 [[arXiv:1007.4716](#)] [[INSPIRE](#)].
- [45] S. Badger, B. Biedermann and P. Uwer, *NGluon: A Package to Calculate One-loop Multi-gluon Amplitudes*, *Comput. Phys. Commun.* **182** (2011) 1674 [[arXiv:1011.2900](#)] [[INSPIRE](#)].

- [46] S. Catani and M.H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [Erratum *ibid.* **510** (1998) 503] [[hep-ph/9605323](#)] [[INSPIRE](#)].
- [47] S. Catani, S. Dittmaier, M.H. Seymour and Z. Trócsányi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189 [[hep-ph/0201036](#)] [[INSPIRE](#)].
- [48] G. Bevilacqua, M. Czakon, M. Kubocz and M. Worek, *Complete Nagy-Soper subtraction for next-to-leading order calculations in QCD*, *JHEP* **10** (2013) 204 [[arXiv:1308.5605](#)] [[INSPIRE](#)].
- [49] Z. Nagy and Z. Trócsányi, *Next-to-leading order calculation of four jet observables in electron positron annihilation*, *Phys. Rev. D* **59** (1999) 014020 [Erratum *ibid.* **62** (2000) 099902] [[hep-ph/9806317](#)] [[INSPIRE](#)].
- [50] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, *Phys. Rev. D* **68** (2003) 094002 [[hep-ph/0307268](#)] [[INSPIRE](#)].
- [51] M. Czakon, H.B. Hartanto, M. Kraus and M. Worek, *Matching the Nagy-Soper parton shower at next-to-leading order*, *JHEP* **06** (2015) 033 [[arXiv:1502.00925](#)] [[INSPIRE](#)].
- [52] J. Alwall et al., *A Standard format for Les Houches event files*, *Comput. Phys. Commun.* **176** (2007) 300 [[hep-ph/0609017](#)] [[INSPIRE](#)].
- [53] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, *Comput. Phys. Commun.* **180** (2009) 2499 [[arXiv:1508.07749](#)] [[INSPIRE](#)].
- [54] Z. Bern et al., *Ntuples for NLO Events at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 1443 [[arXiv:1310.7439](#)] [[INSPIRE](#)].
- [55] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [[arXiv:1706.00428](#)] [[INSPIRE](#)].
- [56] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [[arXiv:1912.10053](#)] [[INSPIRE](#)].
- [57] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204 [[arXiv:1412.3989](#)] [[INSPIRE](#)].
- [58] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [[arXiv:1412.7420](#)] [[INSPIRE](#)].
- [59] M. Jezabek and J.H. Kühn, *QCD Corrections to Semileptonic Decays of Heavy Quarks*, *Nucl. Phys. B* **314** (1989) 1 [[INSPIRE](#)].
- [60] L. Basso, S. Dittmaier, A. Huss and L. Oggero, *Techniques for the treatment of IR divergences in decay processes at NLO and application to the top-quark decay*, *Eur. Phys. J. C* **76** (2016) 56 [[arXiv:1507.04676](#)] [[INSPIRE](#)].
- [61] M. Cacciari, G.P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[INSPIRE](#)].
- [62] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [[arXiv:1506.07443](#)] [[INSPIRE](#)].
- [63] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [[INSPIRE](#)].

- [64] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys. G* **43** (2016) 023001 [[arXiv:1510.03865](#)] [[INSPIRE](#)].
- [65] A. Banfi, G.P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113 [[hep-ph/0601139](#)] [[INSPIRE](#)].
- [66] A. Buckley and C. Pollard, *QCD-aware partonic jet clustering for truth-jet flavour labelling*, *Eur. Phys. J. C* **76** (2016) 71 [[arXiv:1507.00508](#)] [[INSPIRE](#)].
- [67] D. Krohn, M.D. Schwartz, T. Lin and W.J. Waalewijn, *Jet Charge at the LHC*, *Phys. Rev. Lett.* **110** (2013) 212001 [[arXiv:1209.2421](#)] [[INSPIRE](#)].
- [68] *A new tagger for the charge identification of b-jets*, CERN, Geneva (2015) [ATL-PHYS-PUB-2015-040](#).
- [69] ATLAS AND CMS COLLABORATIONS collaboration, *Jet charge determination at the LHC*, CERN, Geneva (2017) [ATL-PHYS-PROC-2017-017](#).
- [70] ATLAS collaboration, *Measurement of the Jet Vertex Charge algorithm performance for identified b-jets in  $t\bar{t}$  events in pp collisions with the ATLAS detector*, CERN, Geneva (2018) [ATLAS-CONF-2018-022](#).
- [71] S. Alekhin, J. Blümlein and S. Moch, *NLO PDFs from the ABMP16 fit*, *Eur. Phys. J. C* **78** (2018) 477 [[arXiv:1803.07537](#)] [[INSPIRE](#)].