Eur. Phys. J. C (2019) 79:915
https://doi.org/10.1140/epjc/s10052-019-7370-7

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Parton showers and matching uncertainties in top quark pair production with Herwig 7

Kyle Cormier[1], **Simon Plätzer**[2], **Christian Reuschle**[3,4] [,a], **Peter Richardson**[5,6], **Stephen Webster**[6]

[1] Department of Physics, University of Toronto, Toronto, Canada
[2] Particle Physics, Faculty of Physics, University of Vienna, Vienna, Austria
[3] Physics Department, Florida State University, Tallahassee, FL 32306-4350, USA
[4] Theoretical Particle Physics, Department of Astronomy and Theoretical Physics, Lund University, 223 62 Lund, Sweden
[5] Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland
[6] IPPP, Department of Physics, Durham University, Durham, UK

Received: 12 December 2018 / Accepted: 6 October 2019 / Published online: 12 November 2019

Abstract We evaluate the theoretical uncertainties in next-to-leading order plus parton shower predictions for top quark pair production and decay in hadronic collisions. Our work is carried out using the HERWIG 7 event generator and presents an in-depth study of variations in matching schemes with two systematically different shower algorithms, the traditional angular-ordered and alternative dipole shower. We also present all of the required extensions of the HERWIG dipole shower algorithm to properly take into account quark mass effects, as well as its ability to perform top quark decays. The predictions are compared at parton level as well as to Large Hadron Collider data, including in the boosted regime. We find that the regions where predictions with a non-top-quark-specific tune differ drastically from data are plagued by large uncertainties which are consistent between our two shower and matching algorithms.

1 Introduction

Top quark pair production is an extremely important process at the Large Hadron Collider (LHC) due to its significant cross section. As the top quark decays before it can hadronize, top quark pair production provides us with a unique opportunity to study Quantum Chromodynamics (QCD) radiation and corrections involving massive particles. This includes measurements of the top quark mass, which is important to constrain the higher-order corrections in the electroweak sector of the Standard Model. Top quark pair production at hadron colliders has become a 'standard candle' due to the accurate calculation [1] and measurement of the total cross section. However, the large production rate also allows the measurement of an ever increasing range of kinematic quantities. This means that different kinematic reconstruction strategies for the top quark, and its mass in particular, including in the boosted regime, can be evaluated. It also means that we can study QCD in detail by comparing with less inclusive calculations and Monte Carlo event generators. The large production rate also means that top quark pair production, particularly with the presence of extra jets from QCD radiation, is often the main background to searches for physics Beyond the Standard Model. A number of measurements are available both extracting the top quark mass [2,3], as well as a number of kinematic properties, e.g. [4].

Monte Carlo event generators [5–7] used for predictions of top quark pair production have seen several improvements, which mainly concentrated on combining next-to-leading order QCD corrections with subsequent parton shower algorithms [8–11], and the production of additional jets using multi-jet merging algorithms, e.g. unitarized schemes [12–15] as employed inside HERWIG 7, or approaches which share a similar spirit [16]. Some of the matching algorithms have addressed off-shell effects in the calculation of the hard process [17], though none of the event generators yet features shower algorithms which properly take into account the effect of the finite width of the top quark and its interplay with the parton shower infrared cutoff.

As compared to indirect approaches based on total cross section measurements, these state-of-the-art simulations provide a very sophisticated description of kinematic properties and thus allow to extract the top quark mass from kinematic properties with an unprecedented precision through template fits. These fits determine the top quark mass parameter used by the event generator simulation. The question in what scheme this mass parameter needs to be interpreted,

[a] e-mail: christian.reuschle@thep.lu.se

and what uncertainties need to be taken into account, is still subject to ongoing research [18–20], and for coherent parton shower evolution in e^+e^- collisions first analytic and numeric insights have been gained on the role of the mass parameter, including measurements of the top quark mass from reconstruction of its decay products [21]. In Ref. [22] the authors consider predictions of observables that are sensitive to the top quark mass, including an evaluation of uncertainties due to scale variations similar to those considered in this paper. Some aspects of the hadronization effects in such observables have also recently been evaluated [23], however a comprehensive analysis of variations in parton shower evolution, and the impact of different paradigms to include higher order corrections has not yet been performed.

The present work is centred around a thorough investigation of how reliable predictions by established paradigms, namely scale variations in matched predictions, are across phase space. This question has not yet been answered by an in-depth comparison of similar, yet algorithmically very different, predictions and their associated variations which can be established to constitute a set of uncertainties when meeting well-defined constraints [24]. We do this particularly in the light of event generator predictions which are highly specialized in their parameter choices and thus might generate a wrong impression of how well theoretical understanding is under control, and the associated question of what improvements, specifically concerning multi-jet merging, are required.

We also use this study to introduce some improvements to both radiation from heavy quarks and the handling of their decays in the HERWIG dipole shower module. These changes enable us to perform this study between different matching and shower algorithms in a consistent way, using the same hadronization and underlying event models and with control over shower starting scales and resummation in the hard emission region.

2 Outline of this work

In this study we use the most recent version, 7.1.4, of the HERWIG event generator to make use of the various improvements to the simulation of heavy quarks in production, shower emissions and decays. The modelling of this physics will be discussed in detail in the following sections. In the version considered HERWIG sets up the next-to-leading order (NLO) QCD corrections to the top quark pair production process using the automated facilities outlined in Ref. [10], using external libraries [25,26] to evaluate the required amplitudes on a phase-space point by phase-space point basis. The production of top quark pairs has been validated against MCFM [27–30]. NLO corrections to the decays are included within a NLO matched parton shower simulation, while we neglect non-factorizable corrections which are beyond the leading contribution in the narrow-width approximation.

Matching the production process to the parton shower is possible within both the subtractive (MC@NLO-type [8]) and the multiplicative (Powheg-type [31]) matching paradigms, using the matching subtractions obtained by the MATCHBOX module along with the QCD corrections required. The matching of the decay to the parton shower is available within the multiplicative paradigm in both the QTilde[1] (angular-ordered) [32] and Dipole (Catani–Seymour [33,34] dipole-type) [35] shower modules.

The details of the simulation inputs including the scale choices, parton distribution functions (PDFs), strong coupling running and analyses used to obtain results are included in the subsequent sections. The default electroweak parameters of HERWIG 7.1 are used in all runs and for the decay corrections the top mass is used as the renormalization scale.

The remainder of this work is organized as follows. In Sect. 3 we consider QCD radiation from the top quark pair production process. In Sect. 4 the parton shower simulation of the decay stage is discussed in detail. We then proceed with an in-depth discussion of the NLO matching in Sect. 5. In Sect. 6 we discuss the parton shower hard scale. We use the framework to assess phenomenologically relevant uncertainties in the matched NLO+PS predictions, and present benchmarks and data comparisons together with a detailed analysis of our findings in Sect. 7. Finally we summarize and give an outlook in Sect. 8.

3 Radiation in production of heavy quarks

3.1 Generalities

For both parton shower algorithms used in the HERWIG event generator, a colour flow is assigned to the hard process on the basis of the tree-level colour sub-amplitudes sq. This is a consequence of evaluating the colour correlations relevant to the soft radiation pattern in the limit of a large number of colour charges, $N_c \to \infty$. The chosen colour flow is used to set the initial conditions in both parton shower modules, in particular identifying which 'dipole'-type systems radiate coherently. Radiation in both parton showers is also subject to a veto on hard emissions, as set by the hard shower scale, to be discussed in more detail in Sect. 7.

Since a comprehensive treatment of non-factorizable QCD effects which connect the production process and the decay beyond the narrow-width-approximation is not avail-

[1] An old-style matrix element correction is used by default in the angular-ordered shower, which is formally equivalent to the Powheg method.

able both parton shower algorithms evolve the production process down to the infrared cutoff which, in the current version, is a cutoff on the relative transverse momentum of the emissions. Once the cutoff has been reached by the evolution of the hard process, the decay of the top quark(s) is performed, and further showering of the decay system is simulated as discussed in Sect. 4.

3.2 Angular-ordered shower

The improved angular-ordered shower used by default in HERWIG is described in detail in Refs. [6,32]. Here we will only summarize the important details relevant for heavy quark production together with recent improvements not described in Refs. [6,32]. The momenta of the partons produced in the parton shower are decomposed in terms of the 4-momentum of the parton initiating the jet, p ($p^2 = m^2$, the *on-shell* parton mass-squared), a light-like reference vector, n, in the direction of the colour partner of the parton initiating the jet and the momentum transverse to the direction of p and n. The four momentum of any parton produced in the evolution of the jet can be decomposed as

$$q_i = \alpha_i p + \beta_i n + q_{\perp i}, \tag{1}$$

where α_i and β_i are coefficients and $q_{\perp i}$ is the transverse four momentum of the parton ($q_{\perp i} \cdot p = q_{\perp i} \cdot n = 0$). If we consider the branching of a final-state parton i to two partons j and k, i.e. $i \to jk$, the evolution variable is

$$\tilde{q}_i^2 = \frac{q_i^2 - m_i^2}{z_i(1 - z_i)}, \tag{2}$$

where q_i^2 is the square of the virtual mass developed by the parton i in the branching, m_i is the physical mass of parton i, and z_i is the momentum fraction of the parton j defined such that

$$\alpha_j = z_i \alpha_i, \quad \alpha_k = (1 - z_i)\alpha_i. \tag{3}$$

The transverse momenta of the partons produced in the branching are

$$q_{\perp j} = z_i q_{\perp i} + k_{\perp i}, \quad q_{\perp k} = (1 - z_i) q_{\perp i} - k_{\perp i}, \tag{4}$$

where $k_{\perp i}$ is the transverse momentum generated in the branching. In this case the virtuality of the parton i is

$$q_i^2 = \frac{p_{Ti}^2}{z(1-z)} + \frac{m_j^2}{z} + \frac{m_k^2}{1-z}, \tag{5}$$

where p_{Ti} is the magnitude of the transverse momentum produced in the branching defined such that $k_{\perp i}^2 = -p_{Ti}^2$.

In this case the probability for a single branching to occur is

$$\mathrm{d}\mathcal{P} = \frac{\mathrm{d}\tilde{q}_i^2}{\tilde{q}_i^2} \frac{\alpha_S}{2\pi} \frac{\mathrm{d}\phi_i}{2\pi} \mathrm{d}z_i P_{i \to jk}(z, \tilde{q}), \tag{6}$$

where $P_{i \to jk}(z, \tilde{q})$ is the quasi-collinear splitting function and ϕ_i is the azimuthal angle of the transverse momentum $k_{\perp i}$ generated in the splitting.

As described in Ref. [32] this choice of evolution variable, including the mass of the radiating parton, together with the use of the quasi-collinear splitting functions gives a better treatment of radiation from the parton in the small-angle region. In this region we expect a suppression of soft radiation for angles $\theta \lesssim m/E$, where θ is the angle of emission, m and E the mass and energy of the radiating parton, respectively. The choices used in HERWIG 7 give the expected smooth turn-off of soft radiation rather than the '*dead-cone*'[2] [36] used in HERWIG 6 [37].

The angular-ordered shower is simulated as a series of individual emissions, and only the shower variables $(\tilde{q}, z, \phi)$ are calculated for each emission. Once the evolution has terminated, i.e. there is no phase space available for further emissions, the external particles are taken to be on-shell and the physical momenta reconstructed.

If we set $\alpha_i = 1$ for final-state progenitors[3] and $\alpha_i = x$, the light-cone momentum fraction, for initial-state progenitors then using Eq. (3) and the momentum conservation relation $\alpha_i = \alpha_j + \alpha_k$, all the α values can be iteratively calculated, starting from the hard process and working outward to the external legs. For final-state radiation the transverse momenta can be calculated in the same way using Eq. (4), whereas for initial-state radiation the transverse momentum is calculated iteratively assuming that the parton extracted from the proton as a result of the backward evolution has zero transverse momentum.[4] The β variables for the external partons can then be calculated using the on-shell condition and those for radiating partons using momentum conservation, i.e. $\beta_i = \beta_j + \beta_k$. The latter step may be iterated until the progenitor is reached giving all the β coefficients.

As a result of the shower evolution all the progenitor partons, I, produced in the hard process gain a virtual mass, i.e. the progenitor partons, which initiated the jets, are no longer on mass shell, $q_I^2 \neq m_I^2$. We need to restore momentum conservation in a way that disturbs the internal structure of the jet as little as possible. The easiest way to achieve this is by boosting each jet along its axis so that their momenta are rescaled, i.e. for every jet a Lorentz boost is applied such that

[2] i.e. radiation was forbidden for $\theta < m/E$.

[3] The partons from the hard process which initiate the parton shower.

[4] Or a non-perturbative '*intrinsic*' transverse momentum.

$$q_I = \left(\mathbf{q}_I; \sqrt{\mathbf{q}_I^2 + q_I^2}\right) \xrightarrow{\text{boost}} q_I' = \left(k_I \mathbf{p}_I; \sqrt{k^2\mathbf{p}_I^2 + q_I^2}\right), \tag{7}$$

where k_I is the rescaling factor. The rescaling factors, and the choice of frame in which to apply the boosts, are determined by the choice of which kinematic variables we wish to preserve in the rescaling process. In Ref. [32] an approach was suggested based on the colour connections between the partons initiating the jets:

- for colour-connected final-state partons the reconstruction was performed in the centre-of-mass frame of the partons and the momenta rescaled such that the centre-of-mass energy was conserved, i.e.

$$\sum_{I=1}^{n} \sqrt{k^2\mathbf{p}_I^2 + q_J^2} = \sqrt{s}, \tag{8}$$

  where $\sqrt{s}$ is the centre-of-mass energy and the same rescaling factor k is used for all the jets;
- for colour-connected initial-state partons the reconstruction is performed in the hadronic centre-of-mass frame and the partonic centre-of-mass energy is preserved. In order to fully specify the kinematics an additional constraint is required which in Ref. [32] was chosen such that the rapidity of the partonic collision was preserved;
- for partons with a colour connection between the initial and final state, such as Deep Inelastic scattering (DIS), the system is reconstructed in the Breit frame of the partons such that the virtuality of the system is preserved.

As the majority of hadronic collisions cannot be decomposed into separate colour-singlet systems in early versions of HERWIG++ hadronic collisions were all reconstructed by first using the procedure for colour-connected initial-state partons and then that for final-state partons. This was changed such that if possible the hard process was decomposed into separate colour-singlet systems,[5] for example in $q\bar{q} \to t\bar{t}$, then the separate colour-singlet systems were reconstructed as described above.

In HERWIG 7 we have adopted an approach which attempts to use as much information as possible on the colour structure of the hard process when performing the reconstruction. In order to achieve this we now consider all the partons in the hard process and commence the reconstruction with the parton which had the hardest, i.e. largest p_T, emission in the parton shower. The system formed by this parton and its colour partner is then reconstructed, with either a full reconstruction of the jet produced by the colour partner, the default, or optionally just using the partner to absorb the recoil leaving it on its partonic mass shell and not performing the reconstruction of the full jet. This procedure is repeated for the parton with the hardest shower emission which has not been reconstructed until all the kinematics of all the jets have been reconstructed. Together with an additional option of preserving the momentum fraction of the softer incoming parton in the hard process, for systems with colour connections between initial-state partons, this means that for a single emission the kinematics reduce to those of the Catani–Seymour [33,34] dipoles making matching in the MC@NLO approach simpler.

[5] In the large number of colours, N_C limit.

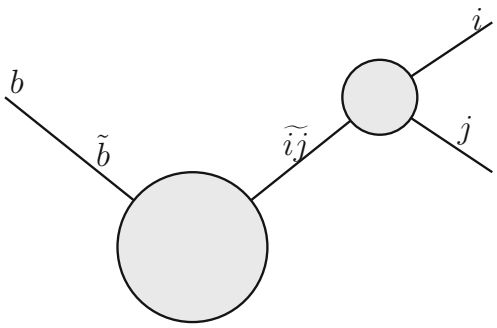
3.3 Dipole shower

The dipole shower algorithm evolves singlet systems of colour connected dipoles, referred to chains [10], based on the colour flow information assigned to the hard process. For the massless case the details of the dipole shower algorithm in HERWIG have been discussed in Refs. [10,35]. In this paper we focus on the generalization of the algorithm to radiation from heavy quarks, and radiation in the decays of coloured objects, to be covered in detail in Sect. 4. While the heavy quark treatment in Ref. [38] has previously been based on Ref. [39], an improved description is presented here which is in one-to-one correspondence to the massless case, and in particular adopts the transverse momentum relevant in the quasi-collinear limit [34], with a smooth massless limit. Throughout this work we use the terminology 'massive dipole' to refer to a dipole that includes at least one massive parton and/or splits to produce at least one massive parton.

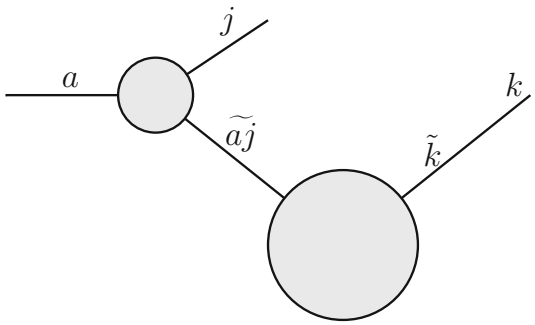
Splittings involving massive incoming partons are not currently implemented in the HERWIG dipole shower. This means that there are three possible dipole configurations involving massive partons. These are shown diagrammatically in Fig. 1. Massive final–final (FF) dipoles, with final-state emitter and spectator, Fig. 1c, must include at least one massive outgoing parton before or after the splitting. Massive final–initial (FI) dipoles, Fig. 1a, consist of a massless incoming spectator and an outgoing emitter. At least one of the outgoing partons before or after the splitting must be massive. Massive initial–final (IF) dipoles, Fig. 1b, consist of a massless incoming emitter and massive final-state spectator.

Due to its large mass, parton shower emissions from top quarks are highly suppressed. This means that emissions from massive FI dipoles do not make a significant contribution in the parton shower. Similarly emissions from FF dipoles with a top quark emitter are highly suppressed, however emissions from FF dipoles with a massless emitter and top quark spectator are not suppressed in this way. Therefore both massive FF and IF dipole splittings make a significant contribution in the parton shower.

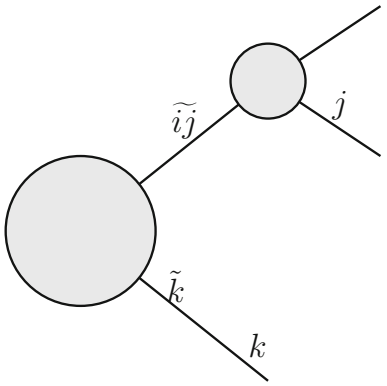
A detailed understanding of these radiation processes with full mass effects is therefore mandatory, and the main goal of



(a) Final-initial dipole.



(b) Initial-final dipole.



(c) Final-final dipole.

Fig. 1 Diagrams of the massive dipoles

this work is to formulate the relevant kinematic parametrization and evolution quantities in a similar way to the massless case, with emphasis on a covariant formulation and an evolution variable which reflects the transverse momentum relevant to the enhancements present for collinear radiation. We present the kinematics used for splittings of all massive dipoles in the following sections. The kernels used to describe the splittings are those given in Ref. [34].

For each dipole the kernels and kinematics used to describe a splitting are parametrized by two splitting variables and an azimuthal angle. In HERWIG 7.1 we use spin-averaged dipole splitting kernels, therefore we randomly generate the azimuthal angle for each splitting according to a uniform distribution. The splitting variables used to parametrize the splitting for each dipole are those used in Ref. [34] and are given for each dipole in the following sections.

In the dipole shower in HERWIG we actually generate the transverse momentum, $p_\perp$, and the light-cone momentum fraction, z, as used in the standard quasi-collinear Sudakov parametrization of the momenta following a splitting. This is the parametrization used in the angular-ordered shower [32], see Sect. 3.2. We choose a light-like vector n to define the collinear direction and for a splitting from a final-state emitter with momentum $\tilde{p}_{ij}$ we write the momentum, q_j, of the emitted parton as

$$q_j = (1-z)\tilde{p}_{ij} + \frac{m_j^2 - (1-z)^2 m_{ij}^2 + p_\perp^2}{2\tilde{p}_{ij}\cdot n(1-z)} n - k_\perp, \tag{9}$$

where m is the mass of the emitted parton and m_{ij} is the mass of the emitter. The space-like vector $k_\perp$ satisfies $k_\perp \cdot \tilde{p}_{ij} = k_\perp \cdot n = 0$ and $k_\perp^2 = -p_\perp^2$.

Similarly for a splitting from a massless incoming parton we write the momentum of the emitted parton as

$$q_j = (1-z)q_a + \frac{p_\perp^2}{2q_a\cdot n(1-z)} n - k_\perp, \tag{10}$$

where q_a is the momentum of the parton incoming from the proton following the splitting and $k_\perp \cdot q_a = k_\perp \cdot n = 0$.

In view of these parametrizations, which are the ones relevant in the (quasi-)collinear limit, we choose to set up kinematic mappings for a dipole splitting including momentum conservation in a way that we express the resulting kinematics in terms of these physical variables, $p_\perp$ and z, rather than the ones which most conveniently allow the separation and integration over the phase space. This has been done in the massless case, and the mappings below generalize this to the massive case with a smooth massless limit.

3.3.1 Final–final dipoles

We consider the splitting process $\tilde{p}_{ij}, \tilde{p}_k \to q_i, q_j, q_k$ where all momenta before and after the splitting are on-shell, $\tilde{p}_{ij}^2 = m_{ij}^2$, $\tilde{p}_k^2 = q_k^2 = m_k^2$, $q_{i,j}^2 = m_{i,j}^2$ and satisfy momentum conservation for the dipole considered, i.e. $Q = \tilde{p}_{ij} + \tilde{p}_k = q_i + q_j + q_k$ with $s = Q^2$. Splittings from FF dipoles are conveniently parametrized by the splitting variables z_i and $y_{ij,k}$ which are defined in terms of the physical momenta as

$$z_i = \frac{q_i \cdot q_k}{(q_i + q_j)\cdot q_k}, \tag{11a}$$

$$y_{ij,k} = \frac{q_i \cdot q_j}{q_i \cdot q_j + q_i \cdot q_k + q_j \cdot q_k}. \tag{11b}$$

A fully consistent mapping from $\tilde{p}_{ij}, \tilde{p}_k \to q_{i,j,k}$ written in terms of z_i and $y_{ij,k}$ is presented in Appendix A.1, however we do not consider it further here. This is because, while this mapping and the corresponding mapping from $q_{i,j,k} \to \tilde{p}_{ij}, \tilde{p}_k$ defined in Ref. [34] are formulated for arbitrary particle masses, the identification of the physical degrees of freedom relevant in the quasi-collinear limit [40] is not directly obvious.

For a massless spectator the relevant direction can be directly identified, however for a massive spectator we first need to map both of the massive dipole momenta prior to emission into light-like momenta n_{ij} and n_k, which in general have $n_{ij} + n_k \neq Q$. We therefore define

$$(n_{ij} + n_k)^2 = 2n_{ij} \cdot n_k \equiv s_{ij,k} \,. \tag{12}$$

We can write these light-like vectors in terms of the emitter and spectator momenta as

$$n_{ij} = \frac{s_{ij,k}^2}{s_{ij,k}^2 - m_{ij}^2 m_k^2} \left(\tilde{p}_{ij} - \frac{m_{ij}^2}{s_{ij,k}} \tilde{p}_k \right), \tag{13a}$$

$$n_k = \frac{s_{ij,k}^2}{s_{ij,k}^2 - m_{ij}^2 m_k^2} \left(\tilde{p}_k - \frac{m_k^2}{s_{ij,k}} \tilde{p}_{ij} \right), \tag{13b}$$

which gives

$$\begin{aligned} s_{ij,k} &= 2n_{ij} \cdot n_k \\ &= \frac{1}{2} \left(s - m_{ij}^2 - m_k^2 + \sqrt{(s - m_{ij}^2 - m_k^2)^2 - 4m_{ij}^2 m_k^2} \right). \end{aligned} \tag{14}$$

The scaled emitter and spectator momenta can be parametrized as

$$\begin{aligned} q_{ij} &= x_{ij} n_{ij} + \frac{m_{ij}^2}{x_{ij} s_{ij,k}} n_k \,, \\ q_k &= x_k n_k + \frac{m_k^2}{x_k s_{ij,k}} n_{ij} \,. \end{aligned} \tag{15}$$

The emitter and spectator momenta relevant in the quasi-collinear limit for the definition of z and $p_\perp$ are expressed as

$$\begin{aligned} q_i &= z q_{ij} + \frac{m_i^2 - z^2 m_{ij}^2 - k_\perp^2}{x_{ij} s_{ij,k} z} n_k + k_\perp \,, \\ q_j &= (1 - z)\, q_{ij} + \frac{m_j^2 - (1 - z)^2 m_{ij}^2 - k_\perp^2}{x_{ij} s_{ij,k} (1 - z)} n_k - k_\perp \,. \end{aligned} \tag{16}$$

Notice that the limit $m_k \to 0$ smoothly reproduces the parametrization where one works with a light-like collinear direction along the spectator. Comparison to Eq. (9) allows us to identify the physical branching variables $p_\perp$ and z, which relate to the propagator involved in the splitting as

$$\begin{aligned} &\frac{1}{z(1-z)} \left(p_\perp^2 + z m_j^2 + (1 - z) m_i^2 - z(1 - z) m_{ij}^2 \right) \\ &\quad = \left[(q_i + q_j)^2 - m_{ij}^2 \right]. \end{aligned} \tag{17}$$

The remaining details of this formulation, including expressions for the scaling variables x_{ij} and x_k and expressions for z_i and $y_{ij,k}$ in terms of the variables $p_\perp$ and z, are provided in Appendix A.2. A formulation similar to that presented here is described in Ref. [41], however it differs in the definition of the momenta of the splitting products and the variables used.

The probability for a single branching to occur from a final–final dipole is

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{(q_i + q_j)^2 - m_{ij}^2} \langle V_{ij,k} \left(z_i, y_{ij,k} \right) \rangle \mathrm{d}q_j \,, \tag{18}$$

where $\langle V_{ij,k} \left(z_i, y_{ij,k} \right) \rangle$ is the spin-averaged dipole splitting kernel used to describe the branching of a final-state emitter into partons i and j with final-state spectator, k. The single-particle emission phase space, discussed in more detail in Appendix A.2, is denoted by $\mathrm{d}q_j$.

Finally, we show that this formulation of the splitting momenta is consistent with the definitions of the kernels and requirements in Ref. [34]. Following the splitting there are three momenta that must be determined, (q_i, q_j, q_k), with no considerations this system contains twelve degrees-of-freedom. Given that we know the identity and therefore the mass of each parton, we can immediately remove three degrees-of-freedom. We are now left with nine degrees-of-freedom, namely the energy, E_n, polar angle, θ_n, and azimuthal angle, ϕ_n, for each parton n, i.e.

$$\begin{aligned} &q_i : \{E_i, \theta_i, \phi_i\}, \\ &q_j : \left\{E_j, \theta_j, \phi_j\right\}, \\ &q_k : \{E_k, \theta_k, \phi_k\}. \end{aligned} \tag{19}$$

We choose to work in the rest frame of the dipole with $\tilde{p}_{ij}$ along the positive z-axis. Implicitly $\tilde{p}_k$ must lie along the negative z-axis and the mapping from $q_{i,j,k} \to \tilde{p}_{ij}, \tilde{p}_k$ defined in Ref. [34] requires that, in this frame, the spectator only absorbs longitudinal momentum in the splitting. Therefore $\theta_k = \phi_k = 0$ which eliminates two degrees-of-freedom. Furthermore we require that the momentum Q is conserved in the splitting which eliminates a further four degrees of freedom. Finally we generate the azimuthal angle of the splitting $\phi = \phi_i = -\phi_j$, where the second equality follows from momentum conservation, according to a uniform distribution. We are now left with two degrees-of-freedom.

It is important to note that the above constraints on the degrees-of-freedom follow from the requirement of momentum conservation in the splitting and the requirements in Ref. [34]. We have also chosen to simplify the picture by working

in a convenient frame which additionally defines the meaning of the azimuthal angle ϕ. Therefore, given ϕ, the momenta following the splitting must be fully constrained by two independent variables. Hence for a given z_i and $y_{ij,k}$ the momenta are fully constrained. Therefore regardless of the variables we generate and the explicit covariant expressions that we use, so long as z_i and $y_{ij,k}$ can be uniquely expressed in terms of the generated variables, the splitting momenta are uniquely defined. Importantly, we can use the splitting kernels and phase-space limits given in Ref. [34] with our covariant formulation of the splitting kinematics.

3.3.2 Final–initial dipoles

As the spectator in a FI dipole is necessarily massless, one can use the standard quasi-collinear parametrization of the kinematics to describe splittings from massive FI dipoles. In order to be consistent with the formulation used to describe splittings from IF dipoles, Sect. 3.3.3, we instead choose to provide a parametrization in terms of the dipole splitting variables. The four-momenta of the spectator and emitter prior to the splitting are $\tilde{p}_b$ and $\tilde{p}_{ij}$, respectively. The four-momenta of the spectator, emitter and emission following the splitting are q_b, q_i and q_j, respectively. The mass of the emitter prior to the splitting and the masses of the emitter and emitted partons following the splitting are m_{ij}, m_i and m_j, respectively.

Splittings from FI dipoles are parametrized by the splitting variables z_i and $x_{ij,b}$ which are defined in terms of the physical momenta as

$$z_i = \frac{q_i \cdot q_b}{(q_i + q_j) \cdot q_b}, \tag{20}$$

$$x_{ij,b} = \frac{(q_i + q_j) \cdot q_b - q_i \cdot q_j + \frac{1}{2}\left(m_{ij}^2 - m_i^2 - m_j^2\right)}{(q_i + q_j) \cdot q_b}. \tag{21}$$

As the spectator is incoming and therefore massless, z_i is identical to the generated variable z. We define the conserved momentum transfer

$$Q = \tilde{p}_{ij} - \tilde{p}_b = q_i + q_j - q_b\,, \tag{22}$$

and for convenience the invariant

$$s_{ij,b} = 2\tilde{p}_{ij} \cdot \tilde{p}_b\,. \tag{23}$$

The momenta prior to the splitting are written in terms of the momenta following the splitting as

$$\tilde{p}_b = x_{ij,b}\, q_b\,, \tag{24}$$

$$\tilde{p}_{ij} = q_i + q_j - (1 - x_{ij,b}) q_b\,. \tag{25}$$

These expressions are satisfied by writing the momenta following the splitting as

$$q_i = z_i \tilde{p}_{ij} + k_\perp + \left[(1 - z_i)\left(\frac{1 - x_{ij,b}}{x_{ij,b}}\right) + \frac{1}{s_{ij,b}}\left(m_i^2 - m_j^2 + (1 - 2z_i)m_{ij}^2\right)\right]\tilde{p}_b\,, \tag{26a}$$

$$q_j = (1 - z_i)\tilde{p}_{ij} - k_\perp + \left[z_i\left(\frac{1 - x_{ij,b}}{x_{ij,b}}\right) + \frac{1}{s_{ij,b}}\left(-m_i^2 + m_j^2 - (1 - 2z_i)m_{ij}^2\right)\right]\tilde{p}_b\,, \tag{26b}$$

$$q_b = \frac{1}{x_{ij,b}}\tilde{p}_b\,. \tag{26c}$$

We obtain an expression for the splitting variable $x_{ij,b}$ in terms of the generated variables $p_\perp$ and z by comparison with Eq. (9), giving

$$x_{ij,b} = \left[1 + \frac{p_\perp^2 + (1 - z)m_i^2 + z m_j^2 - z(1 - z)m_{ij}^2}{s_{ij,b} z (1 - z)}\right]^{-1}. \tag{27}$$

The probability for a single branching to occur from a FI dipole is given by

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{(q_i + q_j)^2 - m_{ij}^2}\frac{1}{x_{ij,b}}\frac{f_b(x_s/x_{ij,b})}{f_b(x_s)} \times \langle V_{ij}^b\left(z_i, x_{ij,b}\right)\rangle \mathrm{d}q_j\,, \tag{28}$$

where $\langle V_{ij}^b\left(z_i, x_{ij,b}\right)\rangle$ is the spin-averaged dipole splitting kernel used to describe the branching of a final-state emitter into the partons i and j with an initial-state spectator, b. The parton density function of the incoming spectator is $f_b(x)$ and x_s is the proton momentum fraction carried by the spectator prior to the splitting, and $\mathrm{d}q_j$ denotes the single-particle emission phase space. A detailed description of the emission phase space is given in Appendix A.3.

3.3.3 Initial–final dipoles

The momenta of the incoming emitter and outgoing spectator prior to the splitting are $\tilde{p}_{aj}$ and $\tilde{p}_k$, respectively. The new emitter following the splitting is defined to be the parton incoming from the proton while the emitted particle is the emitted final-state parton. The momenta of the emitter, emitted particle and spectator following the splitting are q_a, q_j and q_k, respectively. The mass of the spectator is m_k.

Splittings from IF dipoles are parametrized by the splitting variables $x_{jk,a}$ and u_j which are defined in terms of the physical momenta as

$$x_{jk,a} = \frac{q_a \cdot q_j + q_a \cdot q_k - q_j \cdot q_k}{(q_j + q_k) \cdot q_a}\,, \tag{29}$$

$$u_j = \frac{q_a \cdot q_j}{(q_j + q_k) \cdot q_a}\,. \tag{30}$$

We define the conserved momentum transfer

$$Q = \tilde{p}_k - \tilde{p}_{aj} = q_j + q_k - q_a\,, \tag{31}$$

and the invariant

$$s_{aj,k} = 2\tilde{p}_{aj} \cdot \tilde{p}_k\,. \tag{32}$$

The momenta prior to the splitting are written in terms of the momenta following the splitting as

$$\tilde{p}_{aj} = x_{jk,a}\, q_a\,, \tag{33a}$$
$$\tilde{p}_k = q_j + q_k - (1 - x_{jk,a}) q_a\,. \tag{33b}$$

These expressions are satisfied by writing the momenta following the splitting as

$$q_a = \frac{1}{x_{jk,a}}\tilde{p}_{aj}\,, \tag{34a}$$
$$q_j = \left[\left(\frac{1 - x_{jk,a}}{x_{jk,a}}\right)(1 - u_j) - u_j \frac{2m_k^2}{s_{aj,k}}\right]\tilde{p}_{aj} + u_j \tilde{p}_k - k_\perp\,, \tag{34b}$$
$$q_k = \left[\left(\frac{1 - x_{jk,a}}{x_{jk,a}}\right) u_j + u_j \frac{2m_k^2}{s_{aj,k}}\right]\tilde{p}_{aj} + (1 - u_j)\tilde{p}_k + k_\perp\,. \tag{34c}$$

We need to write the splitting variables in terms of the variables generated in the parton shower, $p_\perp$ and z. We set $n = \tilde{p}_k - (m_k^2/s_{aj,k})\tilde{p}_{aj}$ in Eq. (10) and equate this to Eq. (34b) giving

$$x_{jk,a} = \frac{s_{aj,k}}{2r(s_{aj,k} - m_k^2)}(1 - z + r) \times \left[1 - \sqrt{1 - \frac{4r(s_{aj,k} - m_k^2)}{s_{aj,k}} \frac{z(1-z)}{(1 - z + r)^2}}\right], \tag{35a}$$
$$u_j = x_{jk,a}\left(\frac{r}{1 - z}\right), \tag{35b}$$

where we have defined $r = p_\perp^2/s_{aj,k}$. These expressions again relate the backward-evolution, dipole picture recoil to the quantities involved in the physical forward branching process, Eq. 10.

The probability for a single branching to occur from an IF dipole is

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{2q_j \cdot q_a}\frac{1}{x_{jk,a}}\frac{f_a(x_e/x_{jk,a})}{\tilde{f}_{aj}(x_e)} \times \langle V_k^{aj}\left(u_j, x_{jk,a}\right)\rangle \mathrm{d}q_j\,, \tag{36}$$

where $\langle V_k^{aj}\left(u_j, x_{jk,a}\right)\rangle$ is the spin-averaged dipole splitting kernel used to describe the branching of an initial-state emitter $\widetilde{aj}$ into an initial-state parton a and a final-state parton j with a final-state spectator k. The parton density function of the incoming partons $\widetilde{aj}$ and a are $\tilde{f}_{aj}(x)$ and $f_a(x)$, respectively. The proton-momentum fraction carried by the parton $\widetilde{aj}$ is x_e and $\mathrm{d}q_j$ denotes the single-particle emission phase space. A detailed description of the emission phase space is given in Appendix A.4.

4 Radiation in the decays of heavy quarks

In both HERWIG parton showers the production and decay processes are showered independently, following a factorized approach. In the case of top quark pair production, the hard process, e.g. $pp \to t\bar{t}$, is first evolved down to the IR cutoff $\mu_{\text{IR}} \approx 1\text{GeV}$,[6] as described in Sect. 3. This involves radiation from both the initial- and final-state partons, including the top quarks. When simulating predictions with unstable top quarks, these then undergo a perturbative decay, and further shower evolution from the decaying system, and possible further decay products, e.g. those originating form a hadronic W decay. The hard scale relevant for emissions from the decaying top quark is the mass of the top quark, and the evolution will preserve its four-momentum including the virtuality.

Matchbox is currently limited to generating hard processes with on-shell outgoing particles, because in the factorized approach a smearing of the mass with some input distribution consistently is only possible at leading order (LO), and poses major challenges at next-to-leading order unless one resorts to a complete off-shell calculation, which can in principle be handled by the framework. While the angular-ordered shower can handle off-shell coloured particles, the dipole shower can currently only deal with on-shell coloured particles, such that we do not consider a reconstructed resonance hierarchy from a full calculation as an input to the showers. This also implies that in the hard process the top width is set to zero, as we could otherwise not treat it as an on-shell particle at the level of the hard process. In this paper we therefore use a strict narrow-width-approximation, i.e. no mass smearing is applied and the top quark decay width is not included as a scale in the showering.

In HERWIG by default top quarks, t, are decayed according to the 3-body matrix element to a bottom quark, b, and two fermions, f and $\bar{f}'$, via an intermediate W-boson in order that off-shell effects are included for the W-boson. The decay system is then showered as described in Sect. 4.1 for the

[6] A complete physical description, beyond the narrow-width-approximation, would include a separate treatment of radiation above and below the top quark decay width. Such a treatment of particle decays in parton showers is an area for future development in HERWIG.

angular-ordered shower and Sect. 4.2 for the dipole shower, which presents a new development which we cover in detail.

In both cases we first shower the top-bottom-W-boson, tbW, system followed by the W-boson-fermion-antifermion, $Wf\bar{f}'$, system. In the shower the tbW and $Wf\bar{f}'$ systems are considered to be colour isolated from each other and the rest of the process. In this sense each decay system is showered independently from the rest of the process. This pattern of evolving 'down' decay trees, i.e. from the hard process towards the final-state particles, is true for all decays in HERWIG 7.

In both parton showers we have the option of performing the first emission from the decay system according to the real-emission matrix element using the builtin Powheg decay correction [42] for all SM decay processes, including both the top quark and W-boson decay. In practice this is sufficient as the NLO virtual corrections only effect the calculation of the width and not the physical distributions. This is switched on by default in the dipole shower whereas the angular-ordered shower uses a matrix-element correction by default. The dipole shower does not currently include QED radiation. In the case of SM decays involving no coloured particles, for example a leptonic W-boson decay, QED radiation can be generated using the SOPHTY implementation in HERWIG [43] which is switched on by default when the dipole shower is used. The angular-ordered shower does include QED radiation and the SOPHTY implementation is not used in this case.

4.1 Angular ordered shower

The improved angular-ordered shower used in HERWIG proceeds in much the same way for decays as for hard processes. The main difference is the handling of radiation with a coloured decaying particle connected to one of the decay products, e.g. $t \to bW^+$. In order to cover the full soft phase-space region we must have radiation from both the decaying particle and the decay product [32].[7] This can be seen in Fig. 2 where in order to cover the full phase-space region for soft emission, i.e. $x_g \to 0$, we need radiation in both the upper region, from $b \to bg$ branchings, and the lower region from $t \to tg$ branchings. As can be seen in Fig. 2 the shower approximation overestimates the leading-order real emission matrix element over all the filled phase space and the two results agree in the soft $x_g \to 0$ and collinear limit where x_W tens to its maximum value. The angular-ordered shower has a 'dead-zone' where there is no emission from the parton shower, and a region at large x_g which could in principle be filled by the parton shower. In this region the parton shower significantly underestimates the real emission matrix element and therefore as this region contains to soft or collinear enhancements we choose not to generate parton shower emissions in it. As described in detail in Refs. [6,32] the recoil from any shower emissions in this case is absorbed such that any recoil perpendicular to the direction of the W boson in the top rest frame is absorbed by the bottom quark, while the remaining component parallel to the W boson direction is absorbed by the W boson.

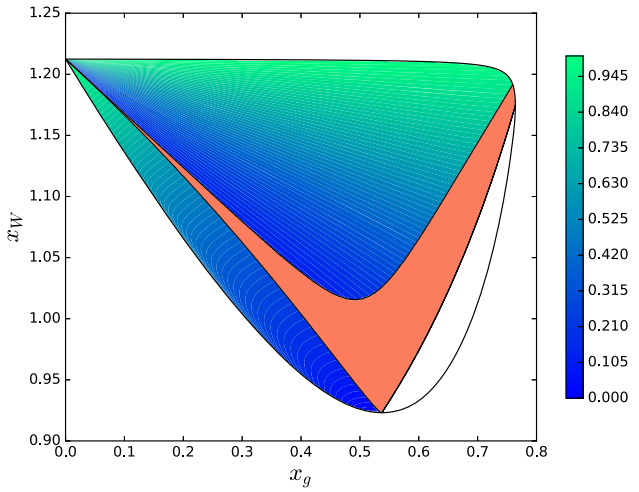


Fig. 2 Dalitz plot for $t \to bW^+g$ where the gluon is emitted by the angular-ordered parton shower. The energy fractions of the gluon and W^+ boson are $x_g = 2E_g/m_{\text{top}}$ and $x_W = 2E_W/m_{\text{top}}$, respectively. In the regions of allowed emission in the angular-ordered parton shower the plot shows the ratio of the leading-order matrix element result over the parton-shower approximation. The red region, the 'dead-zone', is not filled by the parton shower while the empty region for large x_g could be filled by the parton shower, in practice the the shower provides a poor approximation in this region and it and the 'dead-zone' are filled using a hard matrix-element correction

As with HERWIG 6 [44] in HERWIG 7 we apply both a hard matrix-element correction, to fill the 'dead-zone' and unfilled shower region as well as a soft matrix element to correct the hardest-so-far emission in the parton shower regions, this is described in more detail in Ref. [45]. This is the default option, however there is also an option to apply a Powheg correction to the decay [42] including the truncated shower.

4.2 Dipole shower

In a top quark decay a dipole is formed by the decaying top quark and the outgoing bottom quark. During showering the incoming top quark can also form dipoles with other partons outgoing from the decay. In the current implementation of the dipole shower in HERWIG we include emissions from final–initial decay (FI-decay) dipoles only and do not include initial–final decay (IF-decay) dipoles. In other words we explicitly consider emissions from outgoing emitter par-

[7] The original angular-ordered shower in HERWIG 6 [44] did not have radiation from the decaying particle and therefore did not cover the full soft phase-space region.

tons only and do not explicitly include emissions from the incoming top quark. This choice is justified in Sect. 4.2.3.

The dipole splitting kernels for radiation from FI-decay dipoles including only massless final-state particles are given in Refs. [46,47]. In Ref. [48] the dipole splitting kernel for photon radiation from a massive outgoing quark in a FI-decay dipole is presented. The extension to QCD radiation is used to produce the numerical results presented in that paper, however they do not give the explicit form of the splitting kernels used. In all of these works, the authors also decide to include emissions from FI-decay dipoles only.

In this section we describe in detail the treatment of coloured particle decays in the HERWIG 7 dipole shower. The simulation of top quark decays is the primary motivation behind the new developments outlined in this section, therefore we follow the example of top quark decays throughout. These developments have been implemented such that they are applicable to general decays, including BSM processes. In particular the new technical developments in the implementation of the dipole shower, Sect. 4.2.1, and the kinematics for splittings from decay dipoles, Sect. 4.2.2, are independent of the identity of the particles involved.

4.2.1 Implementation

In each decay system the colour chains and dipoles are constructed and updated following each splitting using exactly the same procedure as for the showering of hard production processes [10]. The shower starting scale for each decay system is chosen to be the mass of the incoming decayed particle.

In the case of a top quark decay, with the default POWHEG correction, we attempt to produce the first emission from the tbW system using the exponentiated real-emission matrix element. Following this corrected real emission we shower the system starting from a scale equal to the transverse momentum of the emission. In the rare case that there is no POWHEG emission above the IR cutoff, we do not shower the system.

The $Wf\bar{f}'$ system is a FF dipole, therefore we require no new kinematics or kernels to shower the system. On-the-other-hand the top quark decay introduces new complications. The momentum of the top quark is set, prior to its decay, in the production process and we must not change its momentum following its decay. Therefore in dipoles with the top quark as spectator we cannot use the top quark to absorb the recoil from the splitting. Instead we choose to apply a boost to the rest of the outgoing particles in the decay system to absorb the recoil. This is discussed in more detail in Sect. 4.2.2.

The tbW system is showered until no further emission above the IR cutoff can be generated. This is followed by a 'reshuffling' of the momenta of the particles outgoing from the decay in order to put all partons on their constituent mass-shell as required for hadronization. Momentum conservation is enforced in each splitting in the dipole shower and we must ensure that, following the reshuffling procedure, the sum of the outgoing momenta remains equal to the four-momentum of the decayed particle.

In the case where there are two or more outgoing partons, we simply rescale the masses and 3-momenta of each parton such that all partons are put on their constituent mass shell. In the rare case of no emission from a tbW system, we rescale the mass and 3-momentum of the bottom quark and the 3-momentum of the W-boson while keeping the virtuality of the W-boson unchanged.

Splittings from decay dipoles and the reshuffling procedure can modify the momentum of the W-boson from the value set in the 3-body decay of the top-quark. Therefore following the showering of the tbW system and the subsequent reshuffling, we must apply a boost to the decay products of the W-boson to ensure that momentum is conserved in the W-boson decay. This boost is applied prior to showering the $Wf\bar{f}'$ system. In longer decay trees, following the showering of each decay, we work down the decay tree updating the momenta of decay products as appropriate.

4.2.2 Kinematics

As a colour partner of the emitter we refer to the incoming top quark as the 'spectator', however we wish to preserve the 4-momentum of the top quark as its momentum has been set, before its decay, in the showering of the production process. Therefore the top quark is not used to absorb the recoil in splittings. Instead the recoil is absorbed by all outgoing particles from the top decay system, except for the emitter and the new emission.

Figure 3 shows a diagram of a decay dipole. The momenta of the incoming decayed parton and the outgoing emitter prior to the splitting are q_b and $\tilde{p}_{ij}$, respectively. The total momentum of all other outgoing particles in the decay system is $\tilde{p}_k$. Following the splitting the momenta of the new outgoing emitter and emission are q_i and q_j, respectively and the total momentum of all other outgoing particles in the decay system is q_k. It is implicit from our definition of the recoil system as all particles outgoing from the decay except the emitter that the incoming parton momentum q_b is the conserved dipole momentum

$$Q = q_b = \tilde{p}_{ij} + \tilde{p}_k = q_i + q_j + q_k \,. \tag{37}$$

The splitting kinematics then exactly follow those for a splitting from a massive final–final dipole given in Sect. 3.3.1. The only difference is that for splittings from a decay dipole the recoil, $\tilde{p}_k \to q_k$, is absorbed through the application of an appropriate Lorentz transformation to the recoil system rather than by a single spectator parton. We note that the treatment described here is the same as that used in Ref. [48].

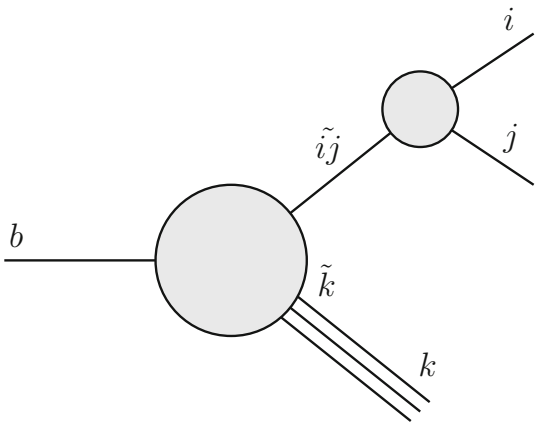


Fig. 3 The final–initial decay dipole

4.2.3 Decay Kernels

As stated in Sect. 4.2 we do not include explicit splittings from IF-decay dipoles. This is because the kernel for a gluon emission from the incoming top quark contains a negative term proportional to the top quark mass-squared which gives rise to a kernel that is almost always negative. We have therefore chosen to include the IF-decay splitting kernels in the FI-decay splitting kernels which are usually large enough to remain positive following the inclusion of the negative mass-squared term. With these considerations there are two possible dipoles and three possible splittings we must consider: the $t-q$ dipole where the final state quark emits a gluon and the $t-g$ dipole where the final state gluon can split into either a $q\bar{q}$-pair or a pair of gluons.

Following the discussion in Sect. 4.2.2 the notation used to express the kernels follows that used for splittings from FF dipoles given in Sect. 3.3.1. We denote the mass of the incoming decayed parton as m_b.

There is only one possible splitting from the $t-q$ FI-decay dipole, $q \to qg$, therefore we must include the entire contribution from the corresponding $t \to tg$ splitting in this kernel. We have used some discretion with regard to which finite pieces are included in the kernels. The kernel, $V_{q\to qg}$, used to describe splittings from a $t-q$ FI-decay dipole in HERWIG 7.1 is given in Eq. (38a). Note that following the conventions of Ref. [34] there is a propagator factor of $1/q_i \cdot q_j$ taken out of the kernel. This is the origin of the factor $y_{ij,k}/(1-z_i(1-y_{ij,k}))$ in front of the $t \to tg$ piece of the kernel and correctly reproduces the eikonal formula that would otherwise be obtained by summing over all possible splittings and configurations for each dipole.

In order to be consistent with the kernels used for splittings from other massive dipoles, we follow the convention from Ref. [34] of multiplying certain terms in the kernels by a finite ratio of relative velocities. The explicit forms of these terms are given in Eq. (39a).

The $t-g$ dipole is more complicated because there are two possible splittings, $g \to gg$ and $g \to q\bar{q}$. The splitting kernels $V_{g\to gg}$ and $V_{g\to q\bar{q}}$ used to describe the $g \to gg$ and $g \to q\bar{q}$ splittings in HERWIG 7.1 are given in Eq. (38b) and Eq. (38c), respectively.

The limits on z_i, $z_{i,\pm}$, and the relative velocity term $v_{ij,i}$ required to express these kernels are given explicitly in Eq. (39c) and Eq. (39d) respectively. We have followed the convention of Ref. [34] and used a parameter κ to distribute finite pieces between the two kernels. In HERWIG 7.1, κ is set to zero in all dipole shower splitting kernels.

We include divergences arising from the IR limits of both q_i and q_j in $V_{g\to gg}$ such that $V_{g\to gg}$ is symmetric with regard to q_i and q_j. This is because this splitting produces indistinguishable final-state gluons and it is consistent with the other $g \to gg$ kernels used in the parton shower.[8]

$$V_{q\to qg} = 8\pi\alpha_S C_F \tag{38a}$$
$$\left\{ \frac{2\left(2m_i^2 + 2y_{ij,k}\bar{s} + \bar{s}\right)}{(1+y_{ij,k})\bar{s} - z_i(1-y_{ij,k})\bar{s}} - \frac{\tilde{v}_{ij,k}}{v_{ij,k}}\left((1+z_i) + \frac{2m_i^2}{y\bar{s}}\right) \right.$$
$$+ \frac{y_{ij,k}}{1-z_i(1-y_{ij,k})}\left[\frac{2\left(2m_i^2 + 2y_{ij,k}\bar{s} + \bar{s}\right)}{(1+y_{ij,k})\bar{s} - z_i(1-y_{ij,k})\bar{s}}\right.$$
$$\left.\left. - \frac{\tilde{v}_{ij,k}}{v_{ij,k}}\left(2 + \frac{2m_b^2}{\left(1-z_i(1-y_{ij,k})\right)\bar{s}}\right)\right]\right\},$$

$$V_{g\to gg} = 16\pi\alpha_S C_A \tag{38b}$$
$$\left\{\frac{1+2y_{ij,k}}{(1+y_{ij,k}) - z_i(1-y_{ij,k})}\right.$$
$$+ \frac{1+2y_{ij,k}}{(1+y_{ij,k}) - (1-z_i)(1-y_{ij,k})}$$
$$\left. + \frac{1}{v_{ij,k}}\left[z_i(1-z_i) - (1-\kappa)z_{i,+}z_{i,-} - 2\right]\right\}$$
$$+ 8\pi\alpha_S C_F$$
$$\left\{ \frac{y_{ij,k}}{1-z_i(1-y_{ij,k})}\left[\frac{2(1+2y_{ij,k})}{(1+y_{ij,k}) - z_i(1-y_{ij,k})}\right.\right.$$
$$\left. - \frac{\tilde{v}_{ij,k}}{v_{ij,k}}\left(2 + \frac{2m_b^2}{(1-z_i(1-y_{ij,k}))\bar{s}}\right)\right]$$
$$+ \frac{y_{ij,k}}{1-(1-z_i)(1-y_{ij,k})}\left[\frac{2(1+2y_{ij,k})}{(1+y_{ij,k}) - (1-z_i)(1-y_{ij,k})}\right.$$
$$\left.\left. - \frac{\tilde{v}_{ij,k}}{v_{ij,k}}\left(2 + \frac{2m_b^2}{(1-(1-z_i)(1-y_{ij,k}))\bar{s}}\right)\right]\right\},$$

$$V_{g\to q\bar{q}} = 8\pi\alpha_S T_R \tag{38c}$$
$$\left\{1 - 2\left(z_i(1-z_i) - (1-\kappa)z_{i,+}z_{i,-} - \frac{\kappa m_i^2}{2m_i^2 + \bar{s}y_{ij,k}}\right)\right\} \tag{38d}$$

[8] The $g \to gg$ splitting can be adjusted to contain only one soft singularity as a means of selecting from the two possible colour flows in that splitting, if a different option has not been pursued, see also the discussions in [49,50].

$$\tilde{v}_{ij,k} = \frac{\sqrt{\lambda\left(s, m_{ij}^2, m_k^2\right)}}{s - m_{ij}^2 - m_k^2}, \tag{39a}$$

$$v_{ij,k} = \frac{\sqrt{\left[2m_k^2 + \left(s - m_i^2 - m_j^2 - m_k^2\right)\left(1 - y_{ij,k}\right)\right]^2 - 4m_k^2}}{\left(s - m_i^2 - m_j^2 - m_k^2\right)\left(1 - y_{ij,k}\right)}, \tag{39b}$$

$$z_{i,\pm} = \frac{2m_i^2 + \left(s - m_i^2 - m_j^2 - m_k^2\right) y_{ij,k}}{2\left[m_i^2 + m_j^2 + \left(s - m_i^2 - m_j^2 - m_k^2\right) y_{ij,k}\right]} \left(1 \pm v_{ij,i} v_{ij,k}\right), \tag{39c}$$

$$v_{ij,i} = \frac{\sqrt{\left(s - m_i^2 - m_j^2 - m_k^2\right)^2 y_{ij,k}^2 - 4m_i^2 m_j^2}}{\left(s - m_i^2 - m_j^2 - m_k^2\right) y_{ij,k} + 2m_i^2}. \tag{39d}$$

Finally we include a symmetry factor of $\frac{1}{2}$, which is not written explicitly here, in front of the $g \to gg$ pieces of $V_{g\to gg}$. With the inclusion of this symmetry factor the factors in front of the eikonal parts from the $g \to gg$ and $q \to qg$ pieces are consistent in the large-N_C limit and we reproduce the correct eikonal expression.

4.2.4 Validation

We present results to validate the new decay kernels and kinematics in the dipole shower. We consider observables which depend primarily on the first, hardest, emission from the decay system and we compare results obtained with and without the real emission decay correction. This comparison directly evaluates how well $V_{q\to qg}$, Eq. (38a), reproduces the full real emission correction. As can be seen in Fig. 4 the kernel overestimates the leading-order matrix element over most of the phase space, apart from a small region near the lower phase-space boundary for $0.1 < x_g < 0.4$.

Our procedure for the following tests exactly follows that used in Refs. [44,45,51]. We generate $e^+e^- \to t\bar{t}$ events at LO at a collision energy of 360 GeV. This collision energy is chosen to be close to the threshold energy for the process, i.e. $2m_{\text{top}}$, in order to reduce radiation from the production process. We work at parton level and include only dileptonic processes. All final-state quarks and gluons are clustered into three jets using the $k_\perp$ algorithm [52] implemented in FastJet [53] and we exclude events containing a jet with transverse energy less than 10 GeV. We additionally exclude events in which the minimum jet separation is less than $\Delta R = 0.7$ where $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$, where η and ϕ denote pseudo-rapidity and azimuthal angle respectively.

We present results for two observables; the separation ΔR_{min} of the closest pair of jets in the event and the jet

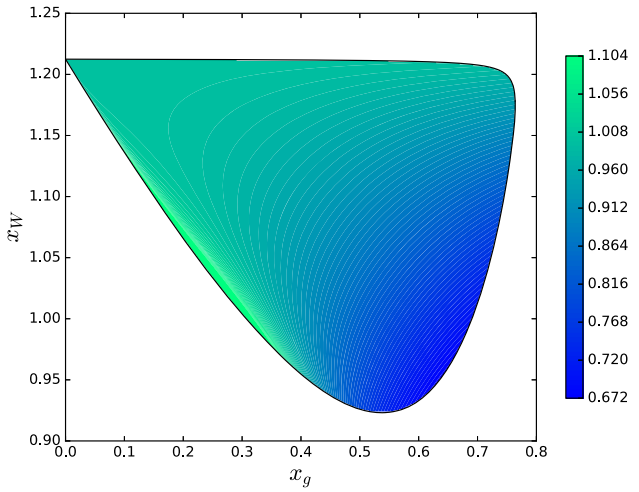


Fig. 4 Dalitz plot for $t \to bW^+g$ where the gluon is showing the ratio of the leading-order matrix element result over the dipole-shower approximation. The energy fractions of the gluon and W^+ boson are $x_g = 2E_g/m_{\text{top}}$ and $x_W = 2E_W/m_{\text{top}}$, respectively

measure y_3, defined as the value of the jet resolution parameter at which the three jet event would be identified as a two jet event. This is given by

$$y_3 = \frac{2}{s}\text{min}_{ij}\left(\text{min}\left(E_i^2, E_j^2\right)\left(1 - \cos\theta_{ij}\right)\right), \tag{40}$$

where s is the centre-of-mass energy squared of the collision, E_i and E_j are the energy of jets i and j respectively and θ_{ij} is the angular separation of jets i and j.

Figure 5 shows the distribution of the minimum jet separation for events showered with and without the real emission decay correction. In general a harder first emission will produce a greater separation of the two closest jets. Therefore, as we expect, the shower with the real emission decay correction produces more events with a larger minimum jet separation. We see that the results with and without the real emission decay correction agree well ($\sim 10\%$) at small jet separations. Furthermore even at large minimum jet separations, where we do not expect the splitting kernel alone to give a good description of the emission, the results agree to within roughly 30%.

Figure 5 also shows the distribution of y_3 for events showered with and without the real emission decay correction. Again, a harder first emission will in general lead to a larger separation of the two closest jets and thus such 2-jet events can be resolved into 3-jet events at a larger value of y_3. As we would expect there is a skew towards larger values of y_3 for the results with the real emission-corrected decay versus the results without the correction. We see that the results at low y_3, corresponding to a softer first emission, are well described by the shower without the real emission decay correction. The log scale used for y_3 in Fig. 5 emphasises the limitations of

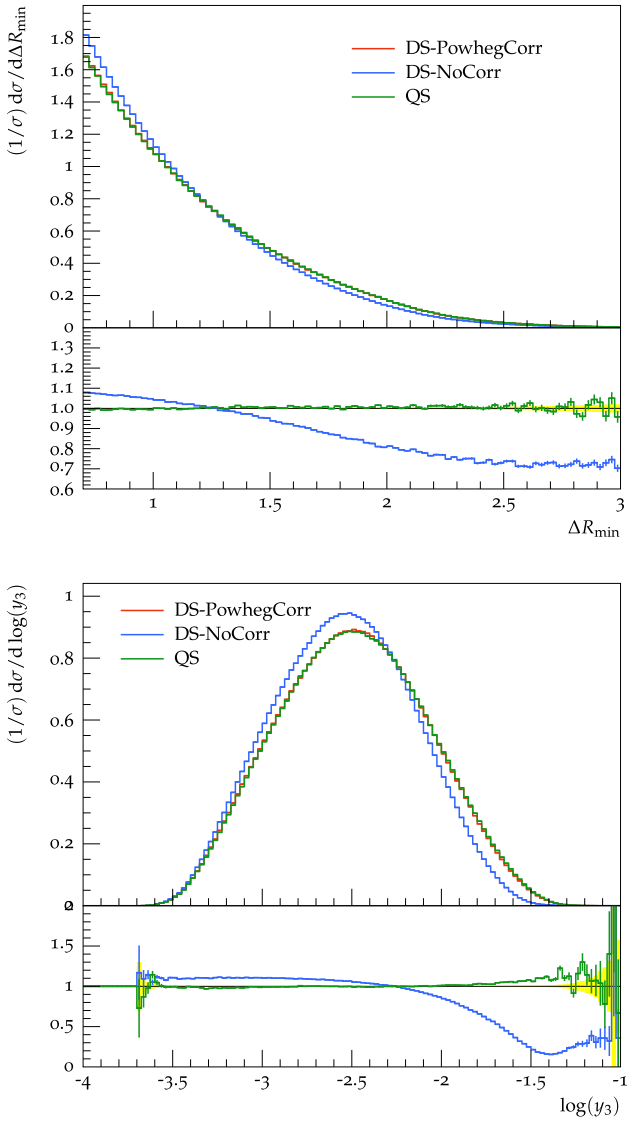


Fig. 5 The distribution of (upper) the minimum jet separation and (lower) the jet measure y_3 in 3-jet $e^+e^- \to t\bar{t}$ events. The distributions are shown for events showered using the dipole shower with (DS-PowhegCorr) and without (DS-NoCorr) the real emission decay correction. In addition we show the distributions obtained using the angular-ordered shower (QS) with the full matrix-element decay correction

the splitting kernel in describing hard emissions. This is evident from the increasing disagreement between the results with and without the real emission decay correction at larger values of y_3.

These results show that the kernel $V_{q\to qg}$ behaves well in the IR region as we require. It also performs reasonably well in the case of harder emissions but its limitations are apparent in the distribution of y_3 in Fig. 5. There is a major limitation to these tests in that they only directly probe the $q \to qg$ splitting kernel. The effects of subsequent emissions are small and it is difficult to create a test to probe $g \to gg$ and $g \to q\bar{q}$ emissions from decay dipoles directly.

As a further comparison we have also included the results from showering with the angular-ordered shower with the appropriate full matrix element correction to the decay in both figures. In all except the lowest bins we see a good agreement between the dipole shower with the real emission decay correction and the angular-ordered shower. This verifies that the corrections in the two showers produce the same behaviour, as we would expect. The disagreement in the lower bins is not a concern as there are numerous differences between the showers and we do not expect agreement to be exact in all regions of phase space.

5 NLO matching and scale choices

A major improvement to the simulation of top quark production and decay in the HERWIG 7 event generator is the inclusion of NLO QCD corrections consistently combined with the subsequent parton shower evolution. NLO matching paradigms are typically less ambiguous than their merging counterparts and entirely driven by solving a matching condition such that the combination of a NLO cross section with a parton-shower evolution reproduces the NLO cross section exactly, plus higher-order terms. In the following we will elaborate on the basic matching algorithms available in HERWIG 7 and their implementation, and will consider in detail the sources of uncertainty involved in matched predictions.

5.1 Hard process setup and NLO subtraction

The partonic cross section for the hard process at leading order can be written as

$$\sigma_{\mathrm{LO}}[u] = \int \mathrm{d}\sigma_B(\phi_n)\mathrm{d}f\; u(\phi_n), \tag{41}$$

where $\mathrm{d}\sigma_B$ is the Born cross section, $\mathrm{d}f$ denotes the partonic luminosity (parton distribution functions), and $u(\phi_n)$ represents a generic observable defined on the Born phase-space point $\phi_n = \{p_a, p_b \to p_1, \ldots, p_n\}$. The HERWIG 7 MATCHBOX module [10] identifies the possible subprocesses contributing to the cross section, and sets up a multi-channel phase-space generator to map the phase-space measure $\mathrm{d}\phi_n$, which includes the momentum conserving δ-function as well as mass-shell constraints.

For a NLO calculation, which we carry out in the dipole subtraction formalism based on Catani–Seymour dipole subtraction [33,34], real emission processes including an additional jet are then identified in the same way as for the leading-order cross section, and the NLO cross section is calculated as

$$\sigma_{\mathrm{NLO}}[u] = \sigma_{\mathrm{LO}}[u] + \sigma_{V+A+C}[u] + \sigma_{R-A}[u]\,, \tag{42}$$

with

$$\sigma_{R-A}[u] = \int \Big[\mathrm{d}\sigma_R(\phi_{n+1}) u(\phi_{n+1}) - \sum_i \mathrm{d}\sigma_A^{(i)}(\phi_{n+1}) u(\Phi_n^{(i)}(\phi_{n+1})) \Big] \mathrm{d}f \,. \tag{43}$$

The first two terms in Eq. (42) contain the leading-order cross section, as well as the finite combination $\sigma_{V+A+C} = \sigma_{V+I+P+K}$ of virtual corrections, analytically integrated subtraction terms, as well as collinear counterterms, which are all defined over the Born phase-space point ϕ_n and handled accordingly. We have further introduced the dipole subtraction terms $\mathrm{d}\sigma_A(\phi_{n+1})^{(i)}$ and the real-emission contributions $\mathrm{d}\sigma_R(\phi_{n+1})$ which are all functions of the real-emission phase-space point ϕ_{n+1}, and the index i runs over the possible dipole configurations, each of which is associated with a particular kinematic mapping $\Phi_n^{(i)}(\phi_{n+1})$ onto the so-called 'tilde' or underlying Born kinematics. The phase-space mappings trigger phase-space convolutions which can be cast into phase-space factorizations upon introducing suitably adapted parton distribution functions

$$\mathrm{d}\phi_{n+1}\mathrm{d}f|_{\phi_{n+1}=\Phi_{n+1}^{(i)}(\phi_n,r)} = \mathcal{J}^{(i)}(\phi_n, r)\mathrm{d}\phi_n \mathrm{d}f^{(i)}\mathrm{d}r, \tag{44}$$

where $\Phi_{n+1}^{(i)}(\phi_n, r)$ is the inverse mapping to the mapping $\Phi_n^{(i)}(\phi_{n+1})$, and r here refers to the collection of variables required to describe the additional emission, i.e. a scale of the emission, a momentum fraction, and an azimuthal variable. We can also associate the respective definitions as functions of the real emission variables, $R^{(i)}(\phi_{n+1})$, such that

$$\Phi_{n+1}^{(i)}(\Phi_n^{(i)}(\phi_{n+1}), R^{(i)}(\phi_{n+1})) = \phi_{n+1}. \tag{45}$$

MATCHBOX uses diagrammatic information to deduce which subtraction terms need to be included, and automatically sets up a cross section in the form above.

5.2 Parton-shower action and matching

The parton-shower action can conveniently be described as

$$\sigma[u] \to \sigma[\mathrm{PS}_{\mu_{\mathrm{IR}}}[u]], \tag{46}$$

where the parton-shower operator up to the first emission is

$$\mathrm{PS}_{\mu_{\mathrm{IR}}}[u](\phi_n) = \prod_i \Delta^{(i)}(\phi_n, \mu_{\mathrm{IR}}) u(\phi_n) + \sum_i \mathrm{d}P^{(i)}(\phi_n, r)\kappa(Q^{(i)}(\phi_n), p_\perp(r))\theta(q(r) - \mu_{\mathrm{IR}}) \times \prod_j \Delta^{(j)}(\phi_n, \mu_{\mathrm{IR}}) u(\Phi_{n+1}^{(i)}(\phi_n, r)). \tag{47}$$

Here $q(r)$ is the evolution variable which we have singled out only in the phase-space limits on the evolution, starting at a hard scale $Q^{(i)}(\phi_n)$ and ending at the infrared cutoff μ_{IR}. The differential splitting probability is the combination of the respective phase-space factors and a ratio of parton luminosities, and the Sudakov form factor starting at the hard configuration is

$$-\ln \Delta^{(i)}(\phi_n, \mu_{\mathrm{IR}}) = \int \mathrm{d}P^{(i)}(\phi_n, r)\kappa(Q^{(i)}(\phi_n), p_\perp(r))\theta(q(r) - \mu_{\mathrm{IR}}). \tag{48}$$

Notice that the constraint on the hard scale is in general not a sharp cutoff, but might be imposed in different ways, see [24] and the discussion below in Sects. 6.1 and 6.2. We have, not accidentally, chosen the same kinematic mapping as has been used for the dipole subtraction terms. Indeed, the kinematic reconstruction algorithm, and not least the kinematics used in the dipole shower and the Powheg correction to be discussed below, resemble, for one emission, exactly the dipole subtraction kinematics, such that we do not need to consider any additional Jacobian factors.

At this point we can expand the shower action to first order in α_S and subtract this contribution from the NLO cross section to set up the matched cross section. To this extent it is worth noting that we can recast both, the integrand of the Sudakov exponent as well as the emission rate multiplied by the Born cross section into another approximate cross section using the inverse of the kinematic mapping,

$$\mathrm{d}\sigma_{\mathrm{PS}}^{(i)}(\phi_{n+1})\mathrm{d}f \equiv \Big[\mathrm{d}\sigma_B \mathrm{d}f^{(i)}\mathrm{d}P^{(i)}(\phi_n, r) \times \kappa(Q^{(i)}(\phi_n), p_\perp(r))\Big]_{\phi_n=\Phi_n^{(i)}(\phi_{n+1}), r=R^{(i)}(\phi_{n+1})} \,. \tag{49}$$

We have explicitly left out the infrared cutoff in this expression for reasons which will soon become clear. The NLO matching subtraction term is then

$$\sigma_{R-A}^{\mathrm{PS}}[u] = \sum_i \int \mathrm{d}\sigma_{\mathrm{PS}}^{(i)}(\phi_{n+1})\mathrm{d}f \times\theta(q^{(i)}(\phi_{n+1}) - \mu_{\mathrm{IR}}) \Big(u(\phi_{n+1}) - u(\Phi_n^{(i)}(\phi_{n+1}))\Big), \tag{50}$$

with the shorthand $q^{(i)}(\phi_{n+1}) = q(R^{(i)}(\phi_{n+1}))$. The NLO matched cross section is

$$\sigma_{\mathrm{NLO}}^{\mathrm{matched}}[u] = \sigma_{\mathrm{NLO}}[u] - \sigma_{R-A}^{\mathrm{PS}}[u] \,, \tag{51}$$

such that $\sigma_{\mathrm{NLO}}^{\mathrm{matched}}[\mathrm{PS}_{\mu_{\mathrm{IR}}}[u]] = \sigma_{\mathrm{NLO}}[u] + \text{h.o.}$ This can be conveniently combined with the dipole book keeping already employed for the fixed-order NLO calculation to yield two contributions to the NLO matched cross section:

$$\sigma_{\mathrm{NLO}}^{\mathrm{matched}}[u] = \sigma_S[u] + \sigma_H[u] \,, \tag{52}$$

with

$$\sigma_S[u] = \sigma_{\mathrm{LO}}[u] + \sigma_{V+I+P+K}[u] + \sum_i \int \Big(\mathrm{d}\sigma^{(i)}_{\mathrm{PS}}(\phi_{n+1})\theta(q^{(i)}(\phi_{n+1}) - \mu_{\mathrm{IR}}) - \mathrm{d}\sigma^{(i)}_{A}(\phi_{n+1}) \Big) \mathrm{d}f\ u(\Phi^{(i)}_n(\phi_{n+1})) \,, \tag{53}$$

which constitutes Born-type configurations, also referred to as $\mathcal{S}$ events, as well as

$$\sigma_H[u] = \int \Big(\mathrm{d}\sigma_R(\phi_{n+1}) - \sum_i \mathrm{d}\sigma^{(i)}_{\mathrm{PS}}(\phi_{n+1})\theta(q^{(i)}(\phi_{n+1}) - \mu_{\mathrm{IR}}) \Big) \mathrm{d}f\ u(\phi_{n+1}), \tag{54}$$

to provide real-emission type configurations, also referred to as $\mathcal{H}$ events. We stress that these contributions cannot yet be used to generate events with finite weights owing to the presence of the infrared cutoff, which allows for configurations with divergent weights, even if the parton-shower approximated cross section would be able to reproduce the full singularity structure of the real emission. Instead, an additional auxiliary cross section

$$\sigma_X[u] = \sum_i \int \mathrm{d}\sigma^{(i)}_X(\phi_{n+1})\mathrm{d}f \times \theta(\mu_{\mathrm{IR}} - q^{(i)}(\phi_{n+1})) \Big(u(\Phi^{(i)}_n(\phi_{n+1})) - u(\phi_{n+1}) \Big) \,, \tag{55}$$

can be added to the matched cross section to eventually yield modified versions of σ_S and σ_H, which can be employed to generate events. In practice, we use the dipole subtraction terms themselves to facilitate this, i.e. $\mathrm{d}\sigma_X = \mathrm{d}\sigma_A$. Note that, for infrared-safe observables u, σ_X only adds power corrections below the infrared cutoff.

5.3 Matching variants

Both the angular-ordered and the dipole showers fit into the framework outlined above, which constitutes the subtractive, or MC@NLO-type, matching in Herwig 7, and the sole task is to determine the shower matching subtraction $\mathrm{d}\sigma^{\mathrm{PS}}_{R-A}$, which we have implemented in a process-independent way in the Matchbox module. These subtractions are indeed very similar to the dipole subtraction terms, but averaged over azimuthal orientation and for colour correlators evaluated in the large-N_c limit. With the recent development of spin-correlation algorithms in both shower modules [54], spin correlations can be restored in these subtractions, and full colour correlations can be justified when using colour matrix-element corrections [50,55], at least for the dipole shower algorithm.

Another choice is a multiplicative, or Powheg-type, matching for which we employ a hardest emission generator, which performs a shower emission using a modified splitting function, or matrix-element correction, determined from the real-emission and Born matrix elements as

$$P^{(i)}(\phi_n, r) \to \frac{w^{(i)}(\Phi^{(i)}_{n+1}(\phi_n, r))}{\sum_j w^{(j)}(\Phi^{(j)}_{n+1}(\phi_n, r))} \times \frac{|\mathcal{M}_R(\Phi^{(i)}_{n+1}(\phi_n, r)|^2}{|\mathcal{M}_B(\phi_n)|^2} \,, \tag{56}$$

for which no complications arise as the full divergent behaviour is reproduced by construction. An additional truncated, vetoed shower needs to be included if the hardest emission generated this way is not the first one to occur. In practice, for the $w^{(i)}$ we use dipole-type partitioned Eikonal factors to perform the weighting into the different singular channels i and use the `ExSample` library [56] to generate emissions according to the Sudakov form factor obtained from the matrix-element correction defined above.

6 Parton shower hard scale

6.1 Profile scale choices

The parton shower hard scale needs to be limited from above in order to avoid the summation of an unphysical tower of logarithms in the Sudakov exponent. To this extent, we have not chosen a fixed starting scale, but a profile scale function $\kappa(Q^{(i)}(\phi_n), p_\perp(r))$ [c.f. Eqs. (47) to (49)]. This function encodes the possibility that not all of the emission phase space should be available to the parton shower. From here on we will generically denote $Q^{(i)}(\phi_n) = Q_\perp$, i.e. we choose the (upper) hard scale $Q^{(i)}(\phi_n)$ manifest as a scale $Q_\perp$ which defines an upper limit on the transverse momentum available to shower emissions.

Several possible parameterizations of the profile scale choices were investigated for leading-order plus parton-shower predictions [24]. We first introduce a hard veto scale $Q_\perp$, which defines an upper limit on the transverse momentum available to shower emissions. By default this is chosen to be the hard process scale, μ_{H}, which in turn is typically set to the factorization and renormalization scale, but may also be chosen independently in Herwig 7. The profile scale choice $\kappa\,(Q_\perp, p_\perp)$ is a function of $Q_\perp$ and the transverse momentum $p_\perp$ of the splitting. For convenience, we define the quantity x as the ratio of these scales

$$x = \frac{p_\perp}{Q_\perp} \,. \tag{57}$$

The default profile scale choice in Herwig 7 is the resummation profile

$$\kappa\left(Q_\perp, p_\perp\right) = \begin{cases} 1 & , \; x \leq 1-2\rho, \\ 1 - \frac{(1-2\rho-x)^2}{2\rho^2} & , \; x \in (1-2\rho,\, 1-\rho], \\ \frac{(1-x)^2}{2\rho^2} & , \; x \in (1-\rho,\, 1], \\ 0 & , \; x > 1, \end{cases} \tag{58}$$

where ρ is a parameter which is set in HERWIG 7.1.4 to $\rho = 0.3$. The resummation profile is one for $p_\perp < (1-2\rho)Q_\perp$, zero for $p_\perp > Q_\perp$, and quadratically interpolates between these regions. It is expected to reproduce the correct towers of logarithms, and switches off the resummation smoothly towards the hard region.

We compare the resummation profile to the hfact profile, which is the damping factor used in POWHEGBOX [57]. The hfact profile is defined as

$$\kappa\left(Q_\perp, p_\perp\right) = \frac{1}{1+x^2}\,. \tag{59}$$

The hfact profile tends to one in the resummation region, i.e. for $p_\perp < Q_\perp$, and to zero in the fixed-order region, i.e. for $p_\perp > Q_\perp$. Unlike the resummation profile the hfact profile does not produce the desired towers of logarithms, as it varies over too broad a range of scales around the hard veto scale. In particular it is not close enough to one in the Sudakov region, i.e. $p_\perp \ll Q_\perp$, and does not enforce a sufficiently effective cutoff on the shower emissions in the hard region, i.e. for $p_\perp \gg Q_\perp$.

In this paper (see Sect. 7.4.1) we investigate some of the impacts of the choice of the profile scale on the prediction of observables using MC@NLO-type matching. For a detailed discussion of the exact properties of the various profile scale choices available in HERWIG 7 we refer the reader to Ref. [24].[9]

[9] As pointed out in Ref. [24] the choice of the profile scale, i.e. how to approach the boundary of hard emissions, is non-trivial and highly relevant in the context of NLO plus parton-shower matching. The choice of the profile scale is essentially constrained by consistency conditions on central predictions (i.e. it should not modify the input distributions of the hard process) and uncertainties (i.e. large uncertainties are expected in unreliable regions or regions where hadronization effects are dominant, as well as stable results are expected in the Sudakov region). It was found in Ref. [24] that the hfact profile does not admit results compatible with these criteria. Instead, using the resummation profile it was found that the angular-ordered and dipole showers are compatible with each other, both in central predictions and uncertainties (despite their very different nature). In addition to studying some of these effects here for top-quark pair production, we would like to point out that choosing a profile scale reminiscent of the resummation profile rather than the hfact profile might also shed some more light on the effects observed in Higgs-boson pair production in Ref. [58].

6.2 Hard veto scale choices

Both shower modules require an upper limit on the transverse momentum of emissions, which is set by a hard veto scale $Q_\perp$ (see previous section). This hard veto scale coincides with the starting scale for the $p_\perp$-ordered dipole shower, and is explicitly implemented as an additional veto for the angular-ordered shower. By default in HERWIG 7, in leading-order events, i.e. Born-type events, $Q_\perp$ is chosen to be the hard process scale μ_H.

For NLO matched predictions, the generated $\mathcal{S}$ and $\mathcal{H}$ events (see Sect. 5.2) separately undergo showering. While $\mathcal{S}$ events constitute Born-type events and are treated as such, several complications arise for $\mathcal{H}$ events.

In MC@NLO-type matching there is no requirement of exact cancellation between the real-emission matrix element and the subtraction term in any region of phase space, as it is possible for the subtracted real-emission cross section still to contain power corrections in the regions where the real emission is soft or collinear. Correspondingly we expect to see a fraction of $\mathcal{H}$ events with a soft and/or collinear emission. In the case of such an $\mathcal{H}$ event it is unnatural to choose the hard veto scale to be of the order of the small transverse momentum of the real emission. Consider for example our case of $t\bar{t}$ production, and say we have an $\mathcal{H}$ event in which the real emission has a transverse momentum of ~ 2 GeV. Given the high energy scales involved in $t\bar{t}$ production, it would be unreasonable to veto all shower emissions with transverse momentum greater than that of the real emission. Instead we need to choose a hard veto scale which is more representative of the scales involved in the process.

In general, as with most scale choices there is no 'correct' choice and we have some freedom in choosing the hard veto scale. By default in HERWIG 7 we choose $Q_\perp = \mu_\mathrm{H}$, for which we typically choose $\mu_\mathrm{H} = \mu_\mathrm{F} = \mu_\mathrm{R}$ with μ_F and μ_R denoting the factorization and renormalization scale respectively. The hard veto scale and the scale of the hard process may also be chosen independently. Overall, given our previous discussion, we desire to choose $Q_\perp$ to be representative of the scales of the objects outgoing from the hard process. In the case of a hard real emission, a hard veto scale that reflects the scale of the real emission should be used. Conversely in the case of a relatively low-scale real emission, a larger scale should be chosen.

Assume for now that we use $Q_\perp = \mu_\mathrm{H}$ and consider an $\mathcal{H}$ event. Common choices for μ_H involve the transverse masses of the top quark and antiquark, often in a linear or quadratic sum. In the case of a very low-$p_\perp$ real emission, the transverse masses of the top quarks will be largely unaffected by the emission. Therefore we would shower such an event from a scale similar to that had there been no emission. Conversely a high-$p_\perp$ real emission on-average increases the sum of the transverse masses of the top quarks, and the presence

of the hard real emission is reflected in the hard veto scale. There are choices for μ_{H} that, while significantly affected by the scale of the real emission, are relatively large over a wide range of real emission scales. If μ_{H} is large enough, the actual maximum scale for showering will be the maximum physically allowed scale, determined from the splitting kinematics, for the first shower emission. In this case, while μ_{H} may be directly affected by the scale of the real emission, the scale of the real emission will have only a small impact on the subsequent showering.

In the case described above one should consider using an alternative choice for $Q_{\perp}$. We have introduced such a scale, which we denote as μ_{a}, in HERWIG 7.1 for use in $t\bar{t}$ production

$$\mu_{\mathrm{a}} = \sqrt{\frac{1}{n_{\mathrm{out}}} \sum_i m_{\perp,i}^2}\,, \tag{60}$$

where n_{out} is the number of particles outgoing from the hard process prior to showering and the sum is over these outgoing particles. This is simply the quadratic mean of the transverse masses of the outgoing particles in the lab frame. In an $\mathcal{H}$ event with a hard real emission, the scale μ_{a} is sensitive to the scale of this real emission. In the case of an $\mathcal{H}$ event with a low-$p_{\perp}$ real emission, μ_{a} is much larger than the scale of the real emission and better reflects the scales in the process. We note that this scale is not smooth in the limit of a soft/collinear emission, i.e. the transition from $\mathcal{H}$ to $\mathcal{S}$ events. In the case of an $\mathcal{H}$ event with a low-$p_{\perp}$ real emission this returns a scale smaller than that in an $\mathcal{S}$ event by a factor $\sqrt{2/3}$ in the soft/collinear limit. We expect the effects of this discontinuity on results to be very small.

In this paper (see Sect. 7.4.2) we investigate some of the impacts of the choice of the hard veto scale on the prediction of observables using MC@NLO-type matching, and how the effects change depending on the choice for the hard process scale. To do this we compare, for each of three different choices for μ_{H}, results obtained using $Q_{\perp} = \mu_{\mathrm{H}}$ and $Q_{\perp} = \mu_{\mathrm{a}}$. The three choices for μ_{H} that we compare are

$$\mu_1 = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{2}, \tag{61a}$$

$$\mu_2 = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{4}, \tag{61b}$$

$$\mu_3 = m_{t\bar{t}}, \tag{61c}$$

where $m_{t\bar{t}}$ is the invariant mass of the $t\bar{t}$-pair.[10]

As always in discussions of scale choices there is no right or wrong choice. The aim of this discussion is to highlight that when we use MC@NLO-type matching we have to make a choice for the hard veto scale. We will show that, depending on the choice for μ_{H}, different choices for $Q_{\perp}$ can have differing and significant effects on our predictions for observables.

[10] We refer the reader to Ref. [59] for a detailed discussion on dynamical scale choices in top-quark pair production.

7 Uncertainty benchmarks and data comparisons

In order to estimate the uncertainty for the event generator predictions we pursue both, variations of the scales involved in the hard production process as well as the scales involved in the subsequent parton showering (see Sect. 7.3). We also consider the impact of different profile scale choices and of different choices for the hard veto scale, dependening on the scale of the hard process (see Sect. 7.4). We consider $t\bar{t}$ pair production in proton-proton (pp) collisions using both 'parton-level', or 'production-level', predictions for stable top quarks and 'particle-level' predictions for unstable top quarks.

Due to the manifold scale choices and variations in our study, we discuss the impact on the observable distributions in some more detail. As mentioned before, in this study we concentrate our efforst in particular on NLO matched predictions. Understanding matched predictions, in the context of systematically disecting the effects of varying all relevant scales as we do in our study, is a relevant prerequisite to understanding similarly systematic uncertainty estimates in more sophisticated setups, like multi-jet merging.[11]

7.1 Production level

All parton-level, or production-level, simulations are done for a centre-of-mass energy of 13 GeV and use the 'benchmark' settings of Ref. [24]. Except for the variations of interest in each section, we use identical input settings for the parton showers and matching schemes in every run. Only QCD radiation is included in the simulations and the same infrared cutoff of $\mu_{\mathrm{IR}} = 1$ GeV (implemented as the minimum transverse momentum cutoff on shower emissions) is used in both showers. We use a mass parameter of $m_t = 174.2$ GeV in the hard process as well as in the subsequent showering algorithms and all other quarks are considered to be massless.

The factorization and renormalization scales are set to the same value $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} \equiv \mu_{\mathrm{H}}$, where our default for the central hard process scale choice is

[11] As the multi-jet merging in HERWIG 7 [12,15] is paved on the matching paradigm in HERWIG 7 (see Sect. 5) it is important to understand the details of the variations and also the corrections to the decay in detail in a matched setup first, which has comparatively less ambiguities. A multi-jet merged study is then a logical follow up, in the context of future work, looking into more exclusive observables, but neither intended in nor within the scope of the current study.

$$\mu_{\mathrm{H}} = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{4}, \tag{62}$$

i.e. half of the average transverse masses of the top and anti-top quarks, unless stated otherwise. This scale choice is motivated by the results of Ref. [59]. We use the default choice, $Q_{\perp} = \mu_{\mathrm{H}}$, for the hard veto scale in all runs apart from those in which this is the scale of interest. Similarly, the resummation profile scale is used in all runs unless otherwise stated.

We use the MMHT2014nlo68cl parton distribution functions (PDFs) along with a two-loop running of α_S with $\alpha_S(M_Z) = 0.12$ both in the parton shower and the hard process.[12] All runs use a four-flavour scheme. All cross sections are rescaled to the NNLO cross section of 815.96 pb [13], calculated using Top++2.0 [60] assuming a top mass of 173.2 GeV and including soft-gluon resummation to next-to-next-to-leading-log order, as are the variations we consider and the envelopes resulting from these variations.

We use a purpose-built analysis written in Rivet [61] to analyse the simulated events. Our analysis considers objects with pseudo-rapidity $|\eta| < 5$, with transverse momentum ordered jets obtained from the anti-$k_{\perp}$ jet algorithm [53,62] with a jet radius of $R = 0.4$.

7.2 Particle level

In contrast to production-level simulations, particle-level simulations include top quark decays, hadronisation and hadronic decays, including tau-lepton decays. Particle-level predictions are used to compare our simulations to experimental data in order to quantify how the different algorithms and their intrinsic uncertainties compare to existing collider data. We use existing and publicly available Rivet analyses, for which the collision energy, $\sqrt{s}$, at which each experimental result was measured and the final-states included are summarised in the following text. Specific details of the experimental analyses are available in the references provided. All of the particle-level measurements presented in this section are taken in the 'combined channel', i.e. including both electron and muon final states. Unless otherwise stated, the hard process scale used to generate these events is

$$\mu_{\mathrm{H}} = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{2}. \tag{63}$$

This scale was chosen because it was found to give rise to reasonable predictions of several observables sensitive to jet activity using MC@NLO-type matching. In particular we compared predictions of several observables included in the publicly available Rivet analyses for Refs. [4,63] obtained using $\mu_{\mathrm{H}} = \mu_{1,2,3}$, i.e. the three scales defined in Sect. 6.2. The resummation profile scale is used in all runs and we use the default choice, $Q_{\perp} = \mu_{\mathrm{H}}$, for the hard veto scale in all runs apart from those in which this is the scale of interest.

The default angular-ordered and dipole shower tunes of Herwig 7.1.1 are used in all particle-level runs with the respective showers. The PDF set used is again MMHT2014nlo68cl while α_S is defined separately by using the tuned value for each shower. We use a five-flavour scheme in the runs using the angular-ordered shower, with massless incoming bottom quarks, and the four-flavour scheme in runs using the dipole shower, which treats partons of a given flavour as having the same mass in both the initial and final states. The masses of the bottom quark and top quark are set to 4.2 GeV and 174.2 GeV, respectively, while all other quarks are considered to be massless.

All distributions that are not normalised to their integral are again scaled to the appropriate next-to-next-to-leading order cross section. The NNLO cross sections are 173.60 pb and 247.74 pb for 7 TeV and 8 TeV collisions, respectively.

7.3 Scale variations

In this section we discuss the parton shower and matching scheme uncertainties that arise from scale variations. We present results for chosen observables that probe various aspects of the simulation and compare with existing data where possible.

Following the approach used in Ref. [24], we estimate the uncertainty on the predictions by considering the variations of three scales:

- the factorization and renormalization scale in the hard process, i.e. the hard process scale $\mu_{\mathrm{H}} = \mu_{\mathrm{R}} = \mu_{\mathrm{F}}$;
- the boundary on the hardness of emissions in the shower, i.e. the hard veto scale $Q_{\perp}$;
- the argument of α_S and the PDFs in the parton shower, i.e. the shower scale μ_{S}.[14]

[12] This refers to an input value which is not used in conjunction with a CMW correction and is only used for the production-level benchmark settings considered here. Typically a tuned value will include the CMW correction numerically. Also, note that in Herwig 7 we perform the running of α_S ourselves rather than using the running determined by the PDF set.

[13] This is the reference cross section calculated by the CMS and ATLAS collaborations.

[14] In this study we are concerned only with variations of the arguments of α_S and the PDFs in the parton showers, therefore, even though they can differ, we use the common terminology 'shower scale' for these scales. In the angular-ordered shower the argument of the strong coupling is related to the transverse momentum of the emitted parton and differs for initial- and final-state evolution, while the argument of the PDFs is simply the ordering variable for initial-state evolution [32]. In the dipole shower the transverse momentum of the emitted parton is used for both scales.

We apply multiplicative factors of 0.5, 1 and 2 to each of the corresponding central scales such that the full set of variations consists of 27 different scale combinations. The complete uncertainty envelope corresponding to this set of variations is shown in each plot. In addition, for each result, we include ratio plots that break down the uncertainties according to the individual scale variations. For each of the three scales considered we separately plot the envelope produced by the upward and downward variations of that scale about the central result, i.e. only two variations are included for each envelope in addition to the central result.

7.3.1 Production level

We first compare results generated with LO matrix elements plus parton shower simulations, using both the angular-ordered (PS) and dipole showers (DS). We use LO plus parton-shower results primarily to compare and contrast the two showers in addition to discussing the uncertainties on the predictions. This is followed by a discussion of results produced by NLO matrix elements matched to a parton shower, i.e. NLO matched simulations. In this discussion, in addition to considering the uncertainties, we focus on the differences between the results obtained using the MC@NLO and Powheg matching schemes.

Figure 6 shows the LO plus parton-shower predictions for the transverse momentum distribution of the top quark ($p_\perp(t)$), the transverse momentum distribution of the $t\bar{t}$-pair ($p_\perp(t\bar{t})$) the jet multiplicity (n_{jets}), and the separation between the $t\bar{t}$-pair and the hardest jet in the event ($\Delta R(t\bar{t}, j_1)$).[15] We find, as expected, that the two shower schemes give similar predictions in the infrared and collinear limits (e.g. at low transverse momentum of the $t\bar{t}$-pair) but can diverge in regions of high momentum emissions. For example the dipole-shower prediction of an increased number of high jet-multiplicity events can be understood by the larger phase space available to the dipole shower due to the fact that there are no angular-ordering restrictions. Nonetheless, in all bins the two predictions agree within the prescribed uncertainties. Furthermore the sizes and sources of uncertainties are similar in all bins for both showers.

The transverse momentum of the $t\bar{t}$-pair, and the number of jets in the event are both sensitive to the hardest emission, and therefore to the cutoff scale used in the shower. The number of jets is also impacted by the strong coupling used in the shower emissions and we see some sensitivity to the shower scale for this observable at high number of jets.

The separation between the $t\bar{t}$-pair and the leading jet is entirely driven by the parton shower since no jet is produced by the LO matrix element. With only one jet the separation must be greater than π, and so the cutoff scale is the dominant source of uncertainty in this region. Separations of less than π are therefore controlled by subsequent jet emissions, which are impacted both by the scale of the hardest jet and the shower scale. Additionally the total uncertainty band, particularly in the region $\Delta R < \pi$, is much larger than the sum in quadrature of the individual components. This demonstrates the fact that the various scales can have correlated effects on the observables which must be taken into account, and suggests the use of the full envelope as done in this paper rather than separate estimation of each of the component followed by combination in quadrature.

Figure 7 shows the NLO-matched predictions for the $p_\perp(t)$ (upper row), and $p_\perp(t\bar{t})$ (lower row) distributions obtained using the angular-ordered (PS, left column) and dipole showers (DS, right column). In an NLO-matched sample these observables are formally predicted with NLO and LO accuracy respectively. Any differences between the the MC@NLO-type (NLO$\oplus$, aka subtractive) and Powheg-type (NLO$\otimes$, aka multiplicative) matching schemes are due to higher-order effects. Accordingly we see good agreement between the matching schemes for both showers. It should be noted that in the highest-$p_\perp$ bin, where uncertainties appear somehwat larger the effect is likely due to statistical uncertainties on the predictions used to construct the uncertainty envelopes.

In the case of $p_\perp(t\bar{t})$ we see much smaller uncertainties in the NLO predictions as compared to the LO predictions in Fig. 6 now that the hardest jet is described by the matrix element calculation. Accordingly, $Q_\perp$ is no longer the dominant uncertainty. Instead the scale of the hard process is relevant for the production of the first jet, playing a significant role across the entire distribution and dominating the uncertainty above $p_\perp(t\bar{t}) \sim 100$ GeV.

Figure 8 shows the NLO-matched predictions of the n_{jets}, and $\Delta R(t\bar{t}, j_1)$ distributions using the angular-ordered and dipole showers. For jet multiplicities above one, and for $\Delta R(t\bar{t}, j_1) < \pi$ these distributions are still controlled largely by the shower. Therefore, they do not benefit from the same magnitude reduction of uncertainties as was seen for $p_\perp(t\bar{t})$. But in both cases μ_{H} now plays a larger role in the overall uncertainty, being relevant for production of the leading jet. The effect of terms beyond NLO accuracy is clearly visible in the jet multiplicity distribution, where the MC@NLO-type matching predicts more high multiplicity events than the Powheg-type matching. This is related to the choice of the hard veto scale as discussed in more depth in Sect. 7.4. The large impact of the hard veto scale on the MC@NLO predictions is clearly seen in the uncertainty band ratio plots. More accurate jet multiplicity distributions can be achieved through the use of multi-jet merging algorithms. However, this is beyond the scope of the paper. Both distributions also

[15] The separation is defined as $\Delta R(t\bar{t}, j_1) = \sqrt{\Delta\phi^2 + \Delta y^2}$, where $\Delta\phi$ and Δy denote the difference in the azimuthal angle and rapidity respectively of the $t\bar{t}$-pair and the hardest jet in the event.

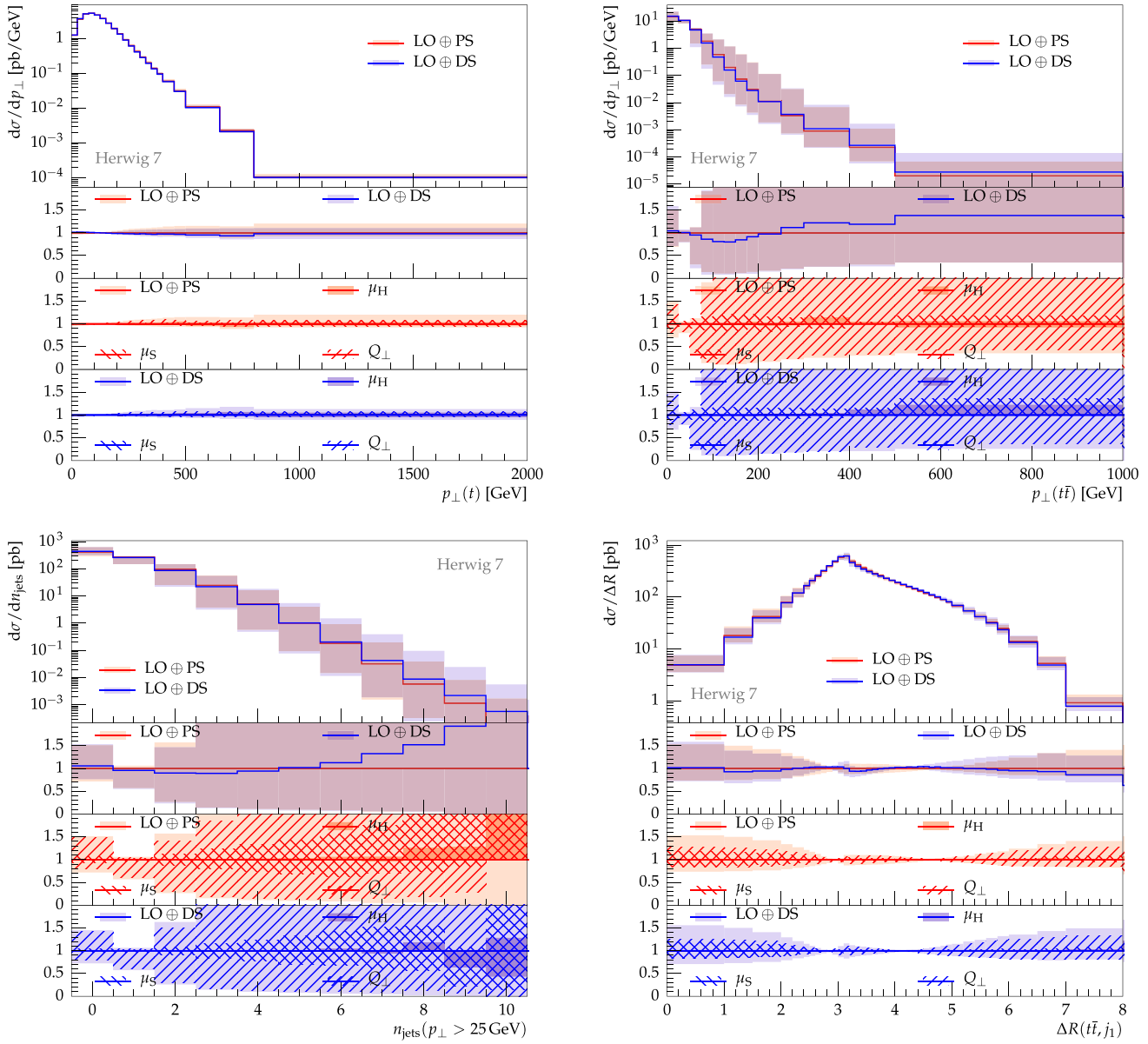


Fig. 6 Scale variations for the inclusive top $p_\perp$-spectrum, the top pair transverse momentum spectrum, inclusive jet multiplicities and R distance between the top pair and the hardest jet using LO plus (LO⊕) parton shower simulations at 13 TeV. In each plot the upper ratio plot compares the envelopes of all variations for the angular-ordered (PS) and dipole (DS) showers, with a ratio to the central prediction of the angular-ordered shower. The bottom two ratio plots in each plot show, for the angular-ordered and dipole showers respectively, a breakdown of all variations into variations of the (factorization and renormalization) scale in the hard process (μ_H), of the arguments of the running coupling and PDFs in the shower (μ_S) and of the hard veto scale ($Q_\perp$)

clealy show the correlated impact of these scale choices with total uncertainty envelopes larger than the sum in quadrature of the individual components.

To summarize this section, for a meaningful uncertainty estimate one must consider which parts of each distribution are well predicted by the hard matrix element and which are filled largely or entirely by the parton shower and one should not expect identical predictions from different parton showers away from the soft and collinear limits. In Monte Carlo studies a thorough evaluation of shower and matching uncertainties is required to account for these differences. Whereas in this section we only considered the production level, in Sect. 7.3.2 we will consider the full particle level, and investigate the uncertainties due to scale variations in the prediction of distributions measured from experiment.

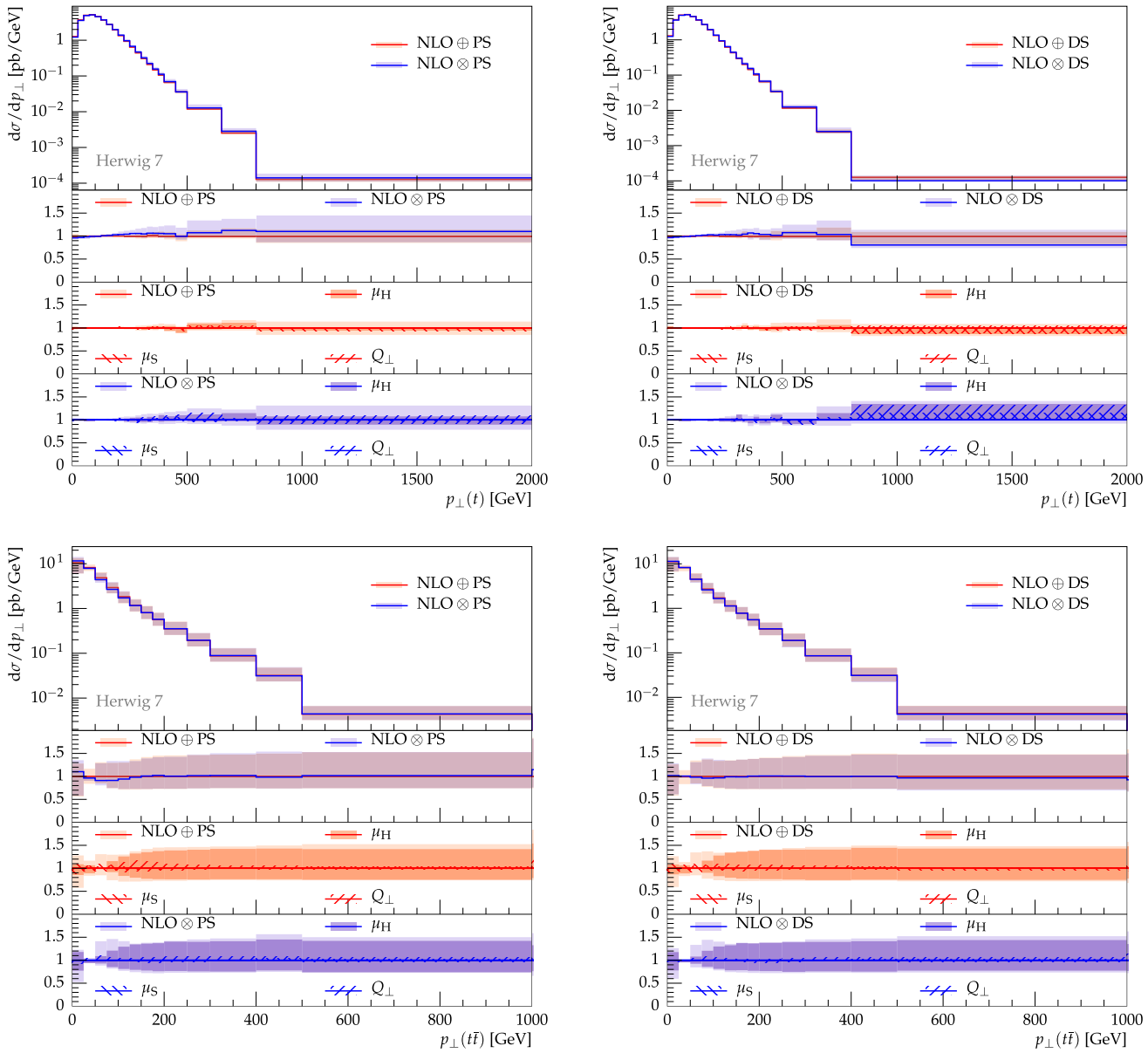


Fig. 7 Transverse momenta of the top quark (upper row) and the top quark pair (lower row), comparing variations for NLO matched predictions at 13 TeV for the angular-ordered (PS, left column) and dipole showers (DS, right column). The top panels in each plot compare the central prediction and overall variation between the MC@NLO-type (NLO⊕) and Powheg-type (NLO⊗) matching. The first ratio plot in each plot allows to directly compare the overall variations in both matching variants, in a ratio to the central MC@NLO-type prediction, while the lower two ratio plots in each plot show a breakdown of the variations for both matching variants regarding the hard process scale(μ_H), the shower scale (μ_S) as well as the hard veto scale ($Q_\perp$)

7.3.2 Particle level

In Sect. 7.3.1 we discussed the uncertainty on predictions of distributions in the production-level process due to scale variations in the simulation. In this section we complete this discussion by looking at the uncertainty on predictions of distributions in the full-process, including top quark decays and hadronization and comparing to available experimental data using publicly available `Rivet` analyses. We perform the same scale variations as in Sect. 7.3.1 and the reader is referred to that discussion for details. We highlight that the veto scale in the showering of decay processes is fixed at the mass of the decayed particle and is not varied.

We examine three observables to compare and contrast with the production-level discussion. Namely we consider the transverse momentum of the hadronic top quark, $p_\perp(t)$,

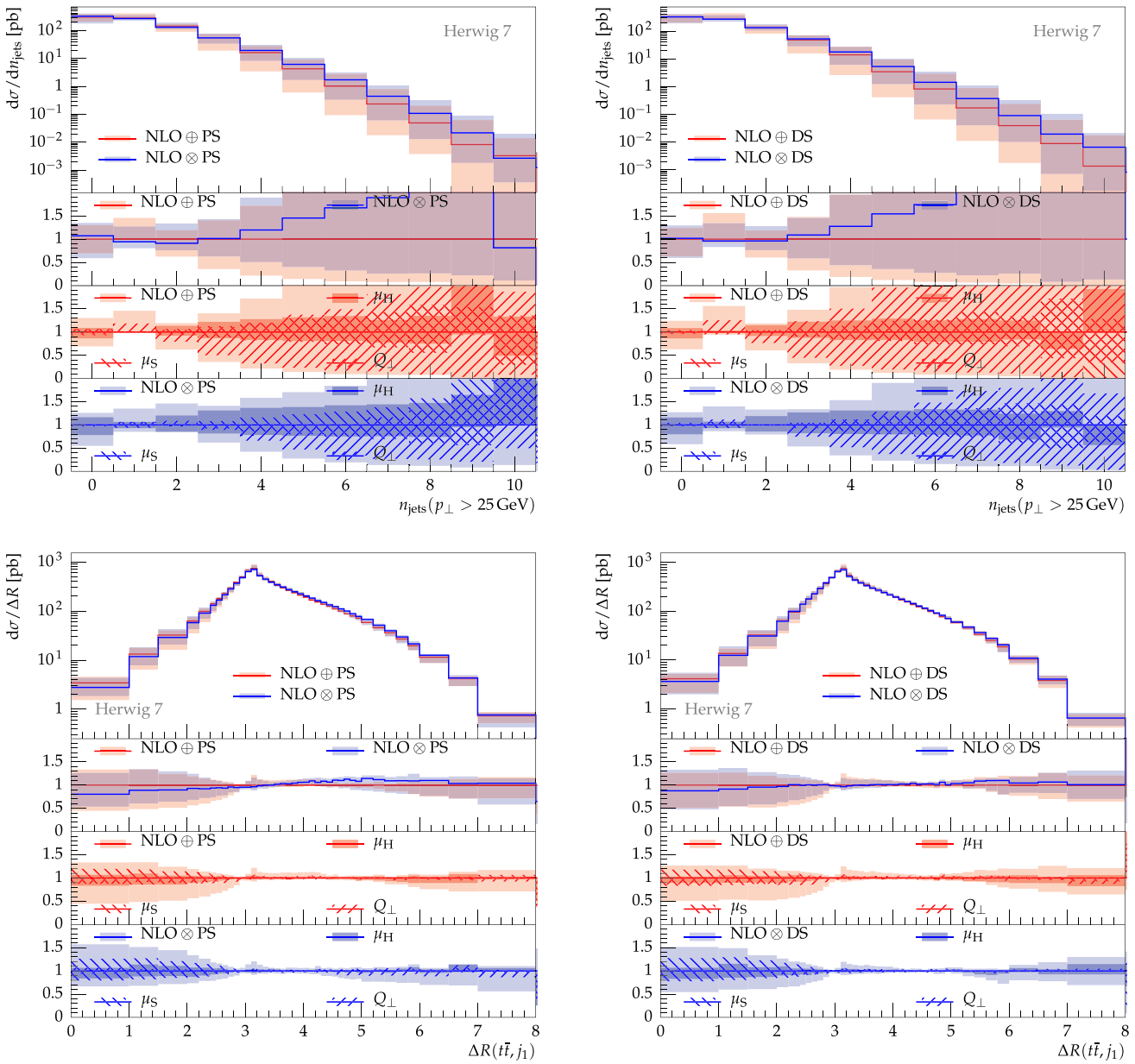


Fig. 8 Same as Fig. 7, in this case showing the inclusive jet multiplicities and the R distance between the $t\bar{t}$-pair pair and the hardest jet. See the text for discussion

the jet multiplicity n_{jets}, and the scalar sum of the transverse momenta H_T (Figs. 9, 10 and 11 respectively). Here, H_T is defined as the sum over all jets in the event.

$$H_{\text{T}} = \sum_{\text{jets}} p_{\perp,\text{jet}} \,. \tag{64}$$

The choice of a different hard process scale [c.f. Eq. (63)] as compared to Sect. 7.3.1 [c.f. Eq. (62)] leads to two noticeable differences in the distributions. Firstly, the LO prediction of the top quark p_T is now sensitive to the shower veto scale variations (comparing Fig. 9 to Figs. 6 and 7). This is because the hard process scale is also used as the central hard veto scale. The increase in this scale allows for the emission of harder jets from the shower which can be hard enough to alter the top p_T distribution. Secondly we note that for this scale choice the number of extra jets in the event predicted by the Powheg-type and MC@NLO-type showers display reasonable agreement for both showers. This is in contrast to the differences between the matching schemes observed in Fig. 8. We have checked that these differences arise from the choice of hard process scale, and not to other differences in the settings used for the full process simulations.

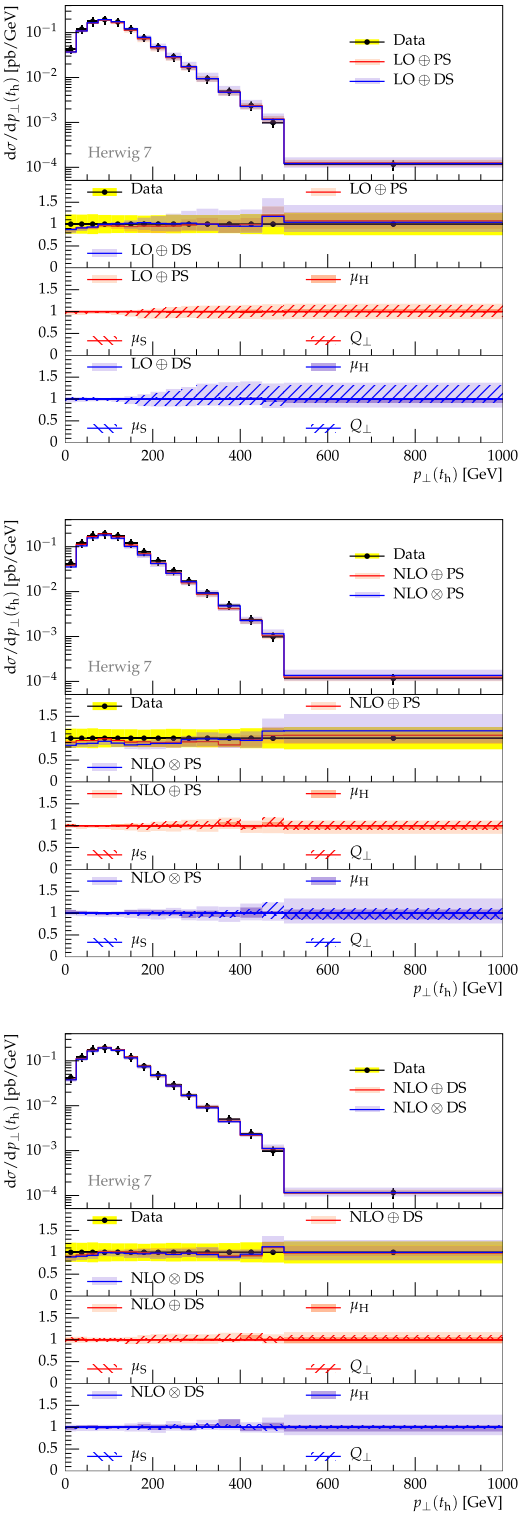


Fig. 9 The transverse momentum of the reconstructed hadronically decaying top quark measured by ATLAS in semileptonic 8 TeV $pp \to t\bar{t}$ events [64]. The top plots shows leading-order production with angular-ordered (PS) and dipole (DS) parton showers, the middle plot NLO production matched to the angular-ordered parton shower while the bottom plot shows NLO production matched to the dipole shower. Two NLO matching schemes, MC@NLO-type (NLO⊕) and Powheg-type (NLO⊗), are used

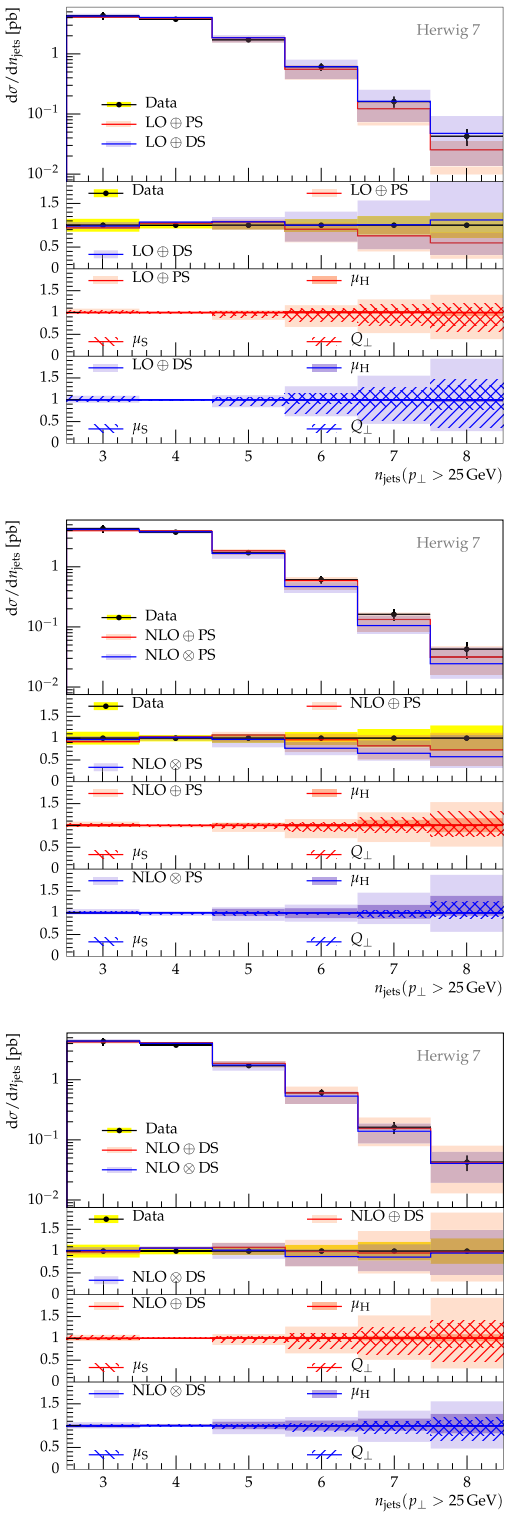


Fig. 10 The multiplicity distribution of jets with transverse momentum greater than 25 GeV, measured in semileptonic 7 TeV $pp \to t\bar{t}$ events by ATLAS [63]. The theoretical predictions are the same as those described in the caption of Fig. 9

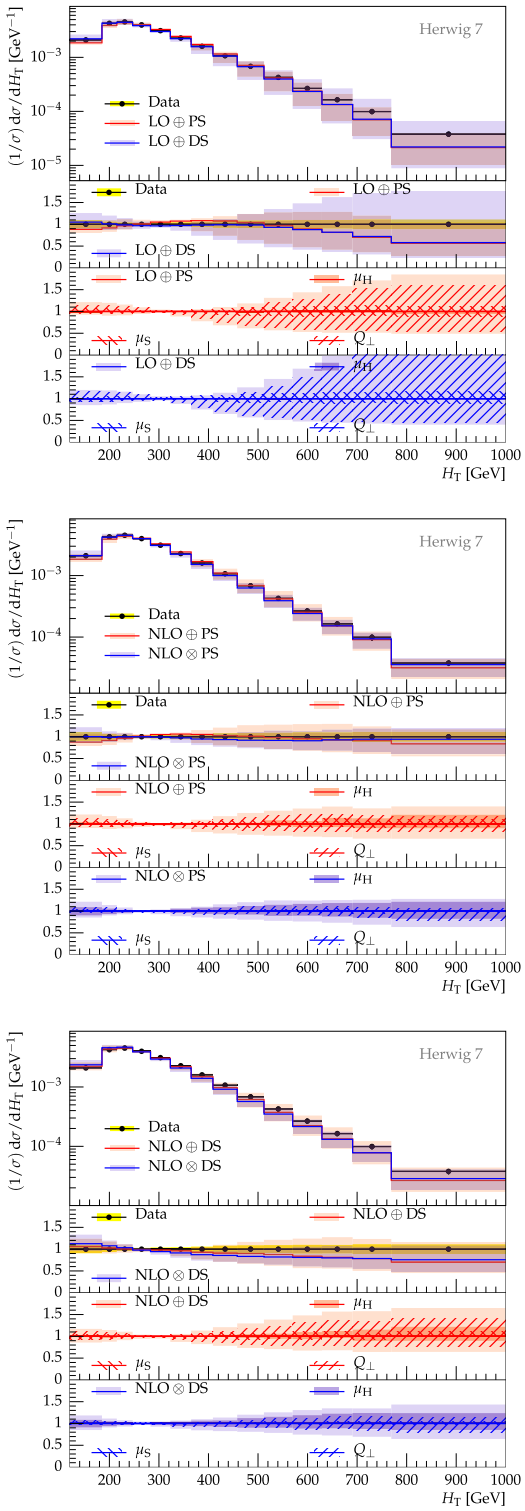


Fig. 11 The H_T distribution measured in semileptonic 8 TeV $pp \to t\bar{t}$ events by CMS [65]. The theoretical predictions are the same as those described in the caption of Fig. 9

As already noted the less restricted phase-space available to the dipole-shower leads to differences between the two algorithms in some regions away from the infrared and collinear limits. In particular the increased n_{jets} prediction as well as the larger sensitivity of the dipole shower to the hard veto scale can be understood as coming from the differences in available phase-space (comparing Fig. 10 to Figs. 6 and 8).

Lastly, we note that the H_T distribution shows sensitivity to both the choice of showering algorith and matching scheme (Fig. 11). Furthermore differences between the Powheg-type and MC@NLO-type matching are pronounced only for the angular shower. This is an example of the correlations between matching scheme and showering algorithm, indicating that the effects cannot be perfectly factorized, similarly to correlations already observed in the scale choices.

7.4 Profile scale and hard veto scale choices

7.4.1 Impact of profile scales

In this study we restrict ourselves to a simple investigation of the effects of the profile scale choice on the simulation of $t\bar{t}$ production using MC@NLO-type matching. To do this we compare results obtained using the two profile scale choices defined above (see Sect. 6.1). This is not intended to be a complete discussion of profile scales and the uncertainties that arise due to choosing a specific one. We simply wish to highlight some of the potential effects of the profile scale choice and present a small selection of observables in which these effects are important.

In Fig. 12 we present production-level results obtained with both the angular-ordered (PS) and dipole showers (DS) using the resummation and hfact profiles. The results shown are the transverse momentum distribution of the hardest jet, $p_\perp(j_1)$, in the upper-left plot, the jet multiplicity, n_{jet}, distributions with a minimum jet-$p_\perp$ cut of 25 GeV and 80 GeV in the upper-right and lower-left plots, respectively, and the distribution of the azimuthal separation of the $t\bar{t}$ pair and the hardest jet, $\Delta\phi(t\bar{t}, j_1)$, in the lower-right plot. For clarity we include a separate ratio plot for each shower which, for each bin, shows the ratio of the result obtained using the hfact profile to the result obtained using the resummation profile.

While the hfact profile suppresses hard shower emissions, it does not apply a hard cut on such emissions as in the resummation profile. We therefore expect to see an increase in both the hardness of the hardest jets and in the number of high-$p_\perp$ jets in events.

Most of the results in Fig. 12 clearly display the expected behaviour, however the jet multiplicity with a minimum jet-$p_\perp$ cut of 25 GeV requires some interpretation. We find that the rate of events with high jet-multiplicities predicted using the hfact profile in the angular-ordered shower is lower than that predicted using the resummation profile. The opposite

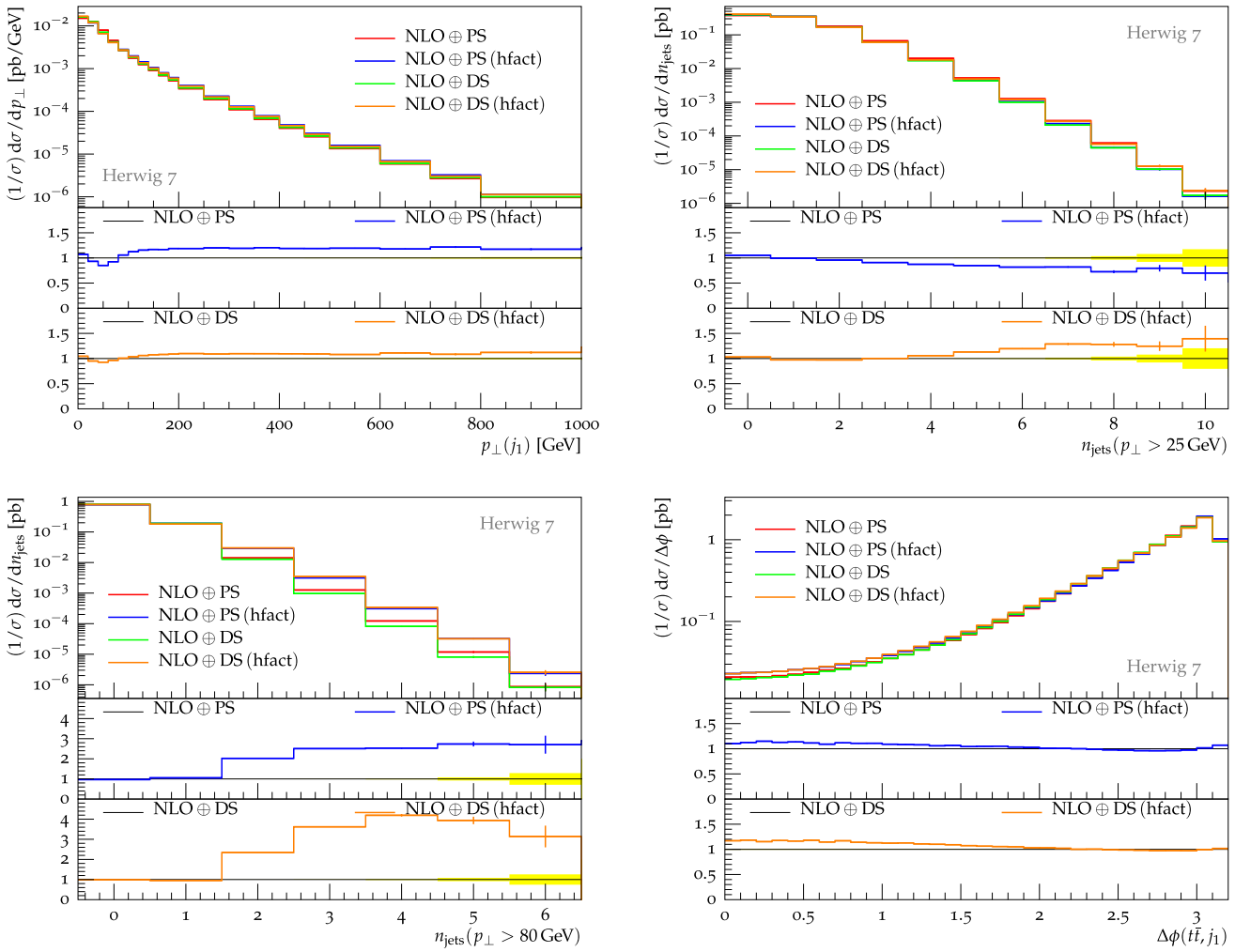


Fig. 12 The effect of different profile scale choices for the two shower algorithms, angular ordered (PS) and dipole (DS), respectively when using MC@NLO-type (NLO⊕) matching. We compare predictions for the default resummation profile versus the broader hfact profile. From left to right, top to bottom, we present the $p_\perp$ spectrum of the hardest jet, the inclusive jet multiplicity at a threshold of 25 GeV and 80 GeV, respectively, as well as the azimuthal angle distance between the top pair and the hardest jet

trend is observed in the results obtained using the dipole shower.

The requirement of angular-ordering in the angular-ordered shower effectively puts a cut on the hardness of shower emissions. Through this the hfact profile can increase the emission phase space only up to a maximum possible value such that the effects of the change from the resummation to hfact profile are expected to be somewhat less pronounced for the angular-ordered shower than the dipole shower. In the case of the angular-ordered shower the angular-ordering restriction and the suppression of soft emissions by the hfact profile lead to a reduction in the number of low-$p_\perp$ jets.

7.4.2 *Impact of hard veto scales*

Next we consider the effect of the choice for the hard veto scale, $Q_\perp$, on results obtained using MC@NLO-type matching. The role of the hard veto scale in MC@NLO-type matching was discussed in Sect. 6.2. In the following we discuss the predictions produced using each of the three options (μ_1, μ_2, μ_3) for μ_{H} separately, the three options being

$$\mu_1 = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{2}, \tag{65a}$$

$$\mu_2 = \frac{m_{\perp,t} + m_{\perp,\bar{t}}}{4}, \tag{65b}$$

$$\mu_3 = m_{t\bar{t}}, \tag{65c}$$

where $m_{t\bar{t}}$ is the invariant mass of the $t\bar{t}$-pair.

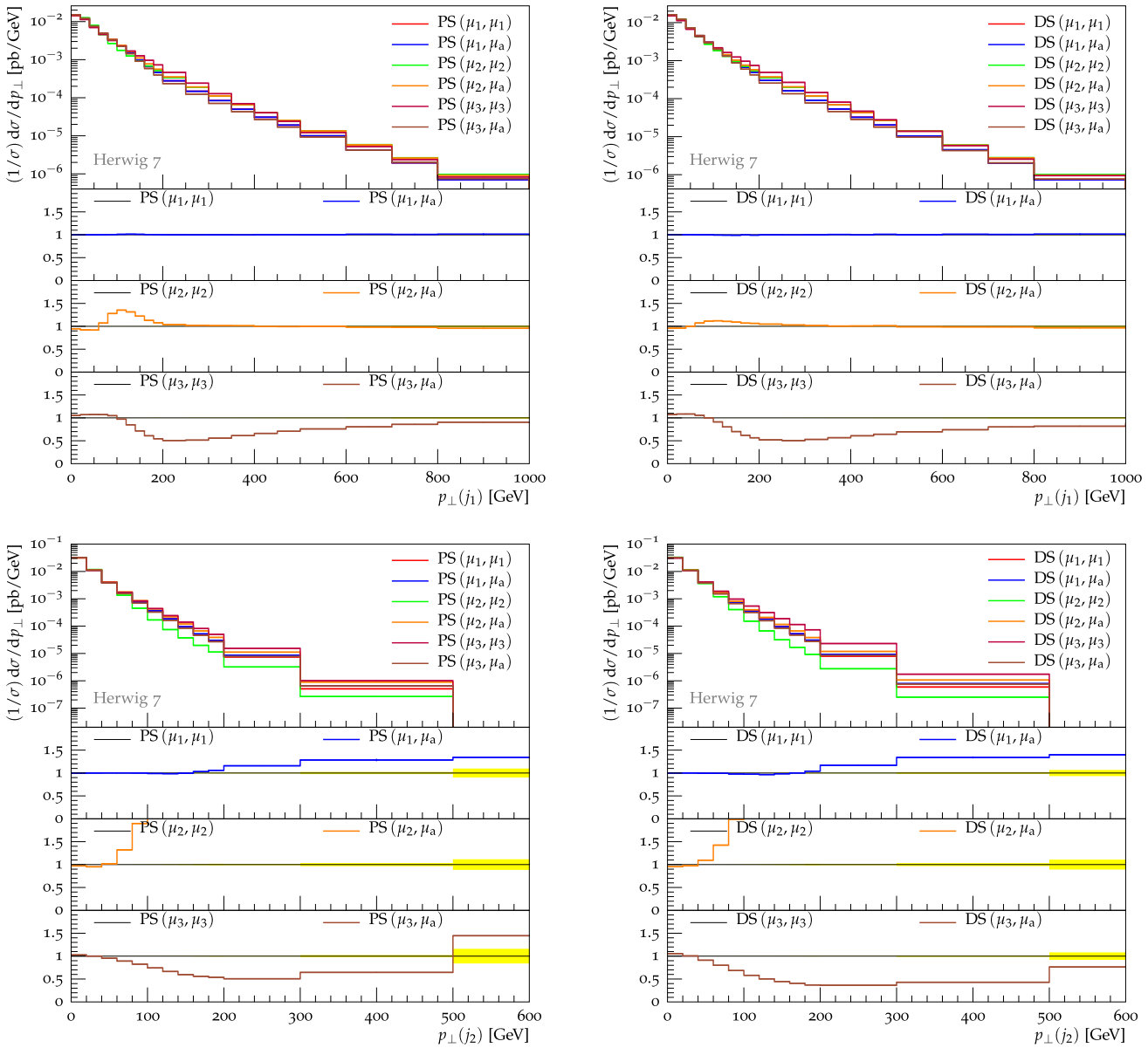


Fig. 13 The effect of different choices of the hard veto scale $Q_\perp$ for the two shower algorithms, angular ordered (PS) and dipole (DS), respectively when using MC@NLO-type matching. We compare predictions for different choices of the resummation and factorization scale choice μ_H, using two choices for the hard veto scale in each case. The scales are specified in the format (μ_H, $Q_\perp$) and each of the scale choices is defined in the text. From top to bottom, left to right, we present the $p_\perp$ spectra of the hardest and second hardest jets, produced with the angular-ordered and dipole shower respectively

Given that $Q_\perp$ directly affects the showering of the production-level process, we expect to see the largest effects due to the choice of $Q_\perp$ (which is either $Q_\perp = \mu_H$ or $Q_\perp = \mu_a$) in distributions that reflect the jet activity in each event. In particular an increase in $Q_\perp$ should increase the jet activity in events.

Figure 13 shows the transverse momentum distributions of the hardest jet, $p_\perp(j_1)$, and second hardest jet, $p_\perp(j_2)$, in production-level events showered using the angular-ordered and dipole showers. These distributions are very sensitive to the hard veto scale as it sets the upper limit on the scale of the first shower emission and affects the available phase space for subsequent emissions. An increase in $Q_\perp$ should produce an increase in the $p_\perp(j_1)$ distribution, however above some scale we expect all distributions to agree regardless of the choice of $Q_\perp$ as the very hardest jets are produced as a NLO real emission in $\mathcal{H}$ events. The scale choices are specified in the format (μ_H, $Q_\perp$).

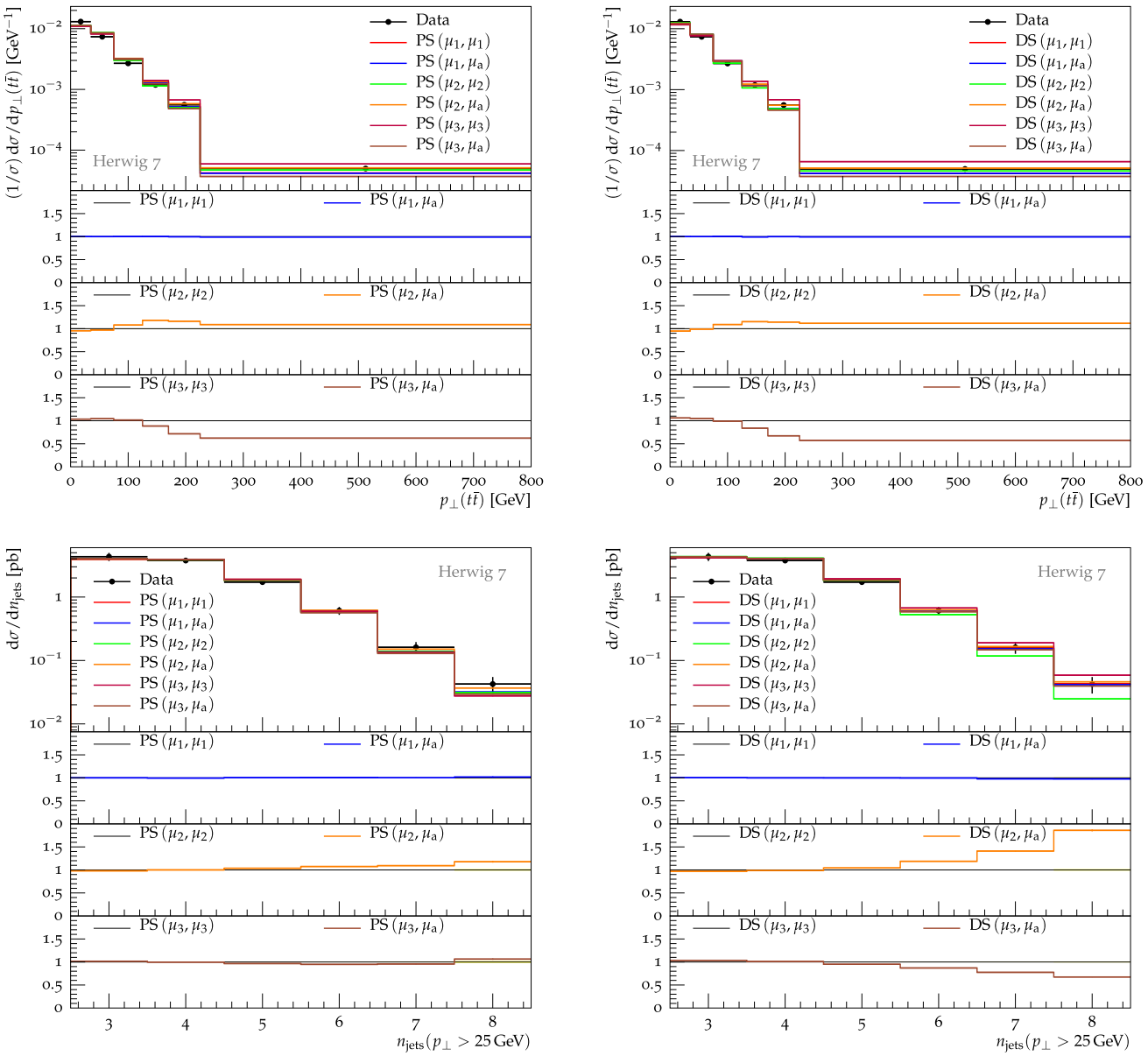


Fig. 14 Top: $t\bar{t}$-pair transverse momentum measured by ATLAS in semileptonic 8 TeV $pp \to t\bar{t}$ events [64] and predicted using the angular-ordered (PS) and dipole (DS) parton showers, respectively. Bottom: The multiplicity distribution of jets with $p_\perp > 25$ GeV measured by ATLAS in semileptonic 7 TeV $pp \to t\bar{t}$ events [63] and predicted using the angular-ordered and dipole parton showers, respectively

In Figs. 14 and 15 we show predictions of several observables measured in particle-level events obtained using the angular-ordered and dipole showers. We highlight that the hard veto scale is applied in the showering of the production process only. The veto scale applied in the showering of the decay process is simply the mass of the decayed particle, i.e. the top quark, and is not varied. We therefore expect the predictions that show the largest change due to the choice of $Q_\perp$ to be those for observables that have a direct dependence on radiation from the production process.

In Fig. 14 the two upper plots show the transverse momentum distribution of the $t\bar{t}$-pair, $p_\perp(t\bar{t})$, measured by ATLAS [64] in semileptonic $t\bar{t}$-events at $\sqrt{s} = 8$ TeV while the two lower plots show the jet-multiplicity distribution measured by ATLAS [63] in semileptonic $t\bar{t}$ events at $\sqrt{s} = 7$ TeV. The $p_\perp(t\bar{t})$ distribution should closely reflect the behaviour observed in the $p_\perp(j_1)$ distribution while we expect an increase in $Q_\perp$ to produce an increase in the rate of events with high jet multiplicity.

The two upper plots in Fig. 15 show predictions of the gap fraction, $f(Q_{sum})$, measured by ATLAS [66] in dileptonic

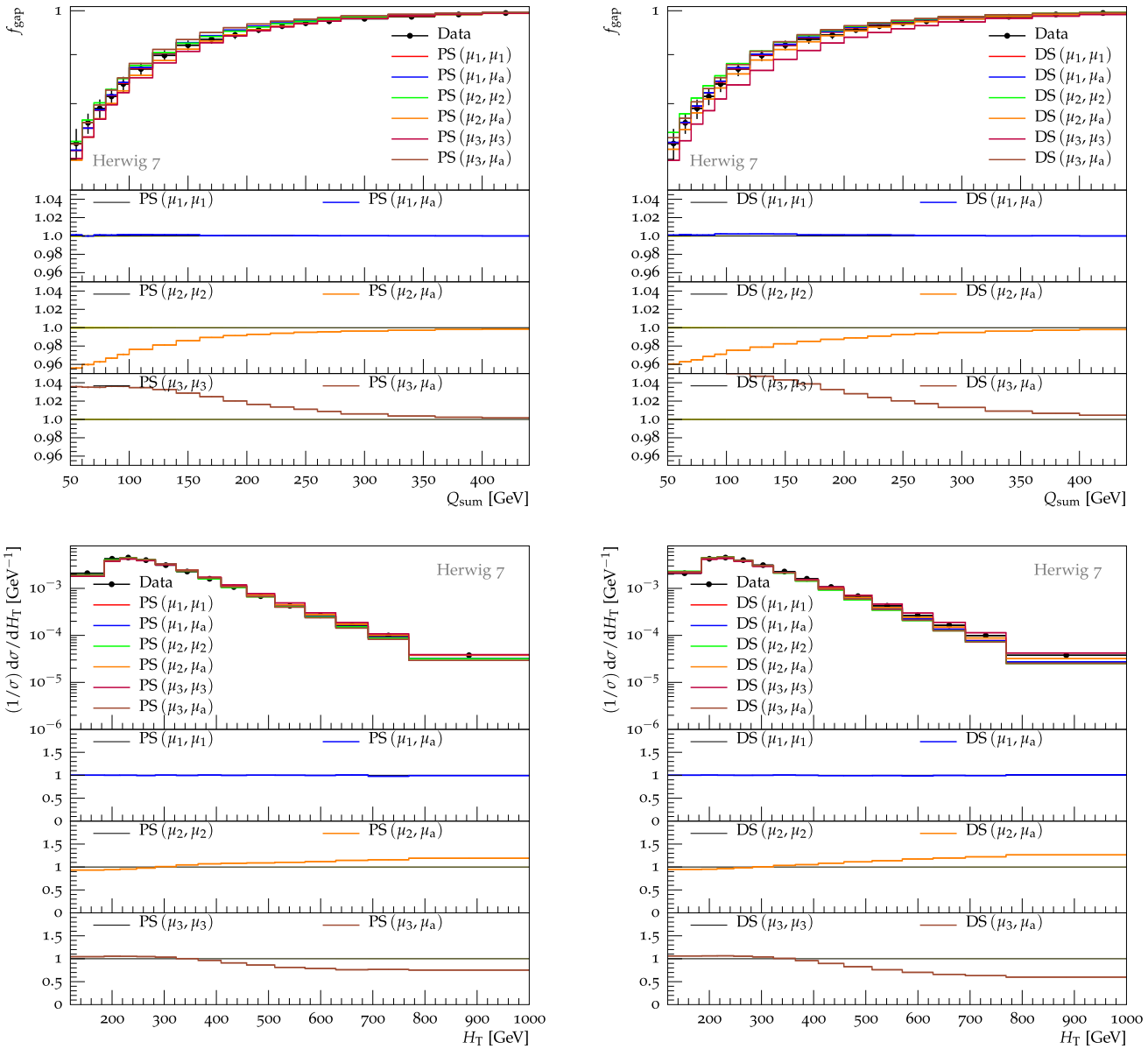


Fig. 15 Top: the gap fraction measured by ATLAS in dileptonic 7 TeV $pp \to t\bar{t}$ events [66], in veto region $|y| < 2.1$, and predicted using the angular-ordered (PS) and dipole (DS) parton showers, respectively. Bottom: combined lepton channel measurement of the H_T distribution by CMS in semileptonic 8 TeV $pp \to t\bar{t}$ events [65] and predicted using the angular-ordered and dipole parton showers, respectively

$t\bar{t}$-events at $\sqrt{s} = 7$ TeV. The gap fraction is a measure of additional jet activity in $t\bar{t}$-events, i.e. jets which originate from quark and gluon radiation in the event as opposed to the decay products themselves. The analysis used selects only events in the dilepton decay channel so that additional jets can be easily distinguished from the decay products, i.e. two leptons and two bottom-tagged jets. The gap fraction is defined as

$$f(Q_{\text{sum}}) = \frac{n(Q_{\text{sum}})}{N}\,, \tag{66}$$

where N is the number of $t\bar{t}$ events that pass the analysis cuts and $n(Q_{\text{sum}})$ is the number of these events in which the sum of the scalar transverse momenta of the additional jets in a given rapidity range is less than the scale Q_{sum}. In particular we present results for additional jets in the rapidity range $|y| < 2.1$. A decrease in the gap fraction corresponds to an increase in jet activity.

Finally, the two lower plots in Fig. 15 show the H_T distribution, as defined in Eq. (64), measured by CMS [65] in semileptonic $t\bar{t}$-events at $\sqrt{s} = 8$ TeV. H_T, which measures

the number and hardness of the jets outgoing from an event, is a sensitive probe of the jet activity in events.

We first consider the choice $\mu_H = \mu_1$ and compare the results for $Q_\perp = \mu_1$ to those for $Q_\perp = \mu_a$. In $\mathcal{S}$-events μ_1 is identical to μ_a and it is only in $\mathcal{H}$-events with the very hardest NLO emissions that μ_a is significantly larger than μ_1. It follows from the similarity of the two scales in most events that most of the distributions in Figs. 13, 14 and 15 indicate very limited difference in the jet activity due to the choice for $Q_\perp$. The change in the $p_\perp(j_2)$ distribution is due to the difference between μ_1 and μ_a in those $\mathcal{H}$-events with the very hardest NLO emissions.

Next we consider $\mu_H = \mu_2$ for which $\mu_a > \mu_H$ in all events. In $\mathcal{S}$-events we have $\mu_a = 2\mu_2$ and in $\mathcal{H}$-events with a low-$p_\perp$ NLO first emission we have $\mu_a \sim \sqrt{8/3}\,\mu_2$. All of the results indicate an increase in the predicted jet activity using $Q_\perp = \mu_a$ compared to using $Q_\perp = \mu_2$ which agrees with our expectation.

As expected the $p_\perp(j_1)$ distribution displays a difference up to some scale, above which the distributions predicted using both choices for $Q_\perp$ come into agreement. The jet multiplicity distribution predicted using the angular-ordered shower displays a more limited difference due to the choice for $Q_\perp$ than the distribution predicted using the dipole shower. This is again due to the additional angular-ordering requirement in the angular-ordered shower.

Finally, we consider the results for $\mu_H = \mu_3$, the invariant mass of the $t\bar{t}$ pair, which is a large scale compared to μ_1 and μ_2 and is larger than μ_a in the vast majority of events. All of the results, other than the jet multiplicity distribution predicted using the angular-ordered shower, indicate a decrease in the jet activity predicted in events using $Q_\perp = \mu_a$ compared to using $Q_\perp = \mu_3$. The difference in the jet multiplicity distribution is again due to the additional angular-ordering requirement in the angular-ordered shower.

To summarize this section, we have compared the effect of using $Q_\perp = \mu_H$ and $Q_\perp = \mu_a$ for three different choices of μ_H. We use μ_a to reflect the transverse momenta of the objects outgoing from the hard process. As there is no first principles choice for the hard veto scale $Q_\perp$, the aim of this discussion is to highlight that when we use MC@NLO-type matching we have to make a choice for this scale. We have demonstrated that the choice of the hard veto scale used in MC@NLO-type matching can have a significant effect on the prediction of observables of interest in $t\bar{t}$ production at the LHC. We have shown that, in general, using a smaller hard veto scale reduces the predicted jet activity in an event, whereas using a larger hard veto scale generally increases the predicted jet activity. We leave further investigation of potential scale choices to future work.

As far as the corrections to the decay and similar variations therein are considered we cannot find any significant impact on the observables considered here, which are mostly insensitive to changes in the decay system.

7.5 Observables sensitive to the decay process

While the production of top quarks is sensitive to the three scales (μ_H, μ_S, $Q_\perp$) investigated in this paper, the decay of a given top quark is only directly impacted by the choice of shower scale. The other scales are fixed at the mass of the decaying particle for the decay process. While there is always some interplay between the production and decay process we consider here observables sensitive to the decay process. For these observables we use the same settings as for comparisons to data, including the full process simulation. First we consider jet substructure in boosted tops, followed by an examination of the separation of b-jets in ttbar events.

The energy and luminosity provided at the LHC allow studies of top quarks with transverse momenta much higher than the top mass. In such cases the decay products of the top quark are generally not well separated. The b quark, and decay products from the W boson are often collimated, forming a single large jet referred to as a 'boosted' top jet. This topology has several distinct difficulties compared to the lower momentum cases.

Firstly, large-radius jets originating from top quarks need to be discriminated from large-radius jets originating from other coloured particles or from the decays of W and Z bosons. This discrimination, referred to as tagging, typically makes use of the substructure of the large jet. The three pronged nature of the top-quark decay leaves a characteristically three-pronged structure within the large jet which is not usually found in boson decays or pure QCD jets. In practice many different techniques are used to tag large jets as originating from a top quark. Whether it is through machine learning applied directly to jet-algorithm inputs or techniques based directly on high-level observables designed to provide substructure information, these taggers all ultimately make use of the distribution of energy within a jet to perform tagging. The performance of taggers is often estimated from simulation and it is therefore important to understand the impact of the various choices made in the Monte Carlo simulation on the description of the substructure of large jets originating from boosted top quarks.

As a probe of the sensitivity of jet substructure to the Monte Carlo approach we examined the N-subjettiness [67] of boosted top quarks produced with HERWIG 7. N-subjettiness measures the degree to which the constituents of a subjet are collimated along its N primary axes. Ratios of N-subjettiness values for different values of N are often used to tag large-radius jets. The ratios of 2-subjettiness to 1-subjettiness (τ_{21}) and 3-subjettiness to 2-subjettiness (τ_{32}) were compared using different HERWIG 7 settings as shown in Fig. 16. Variation of the shower scale, μ_s, is found to

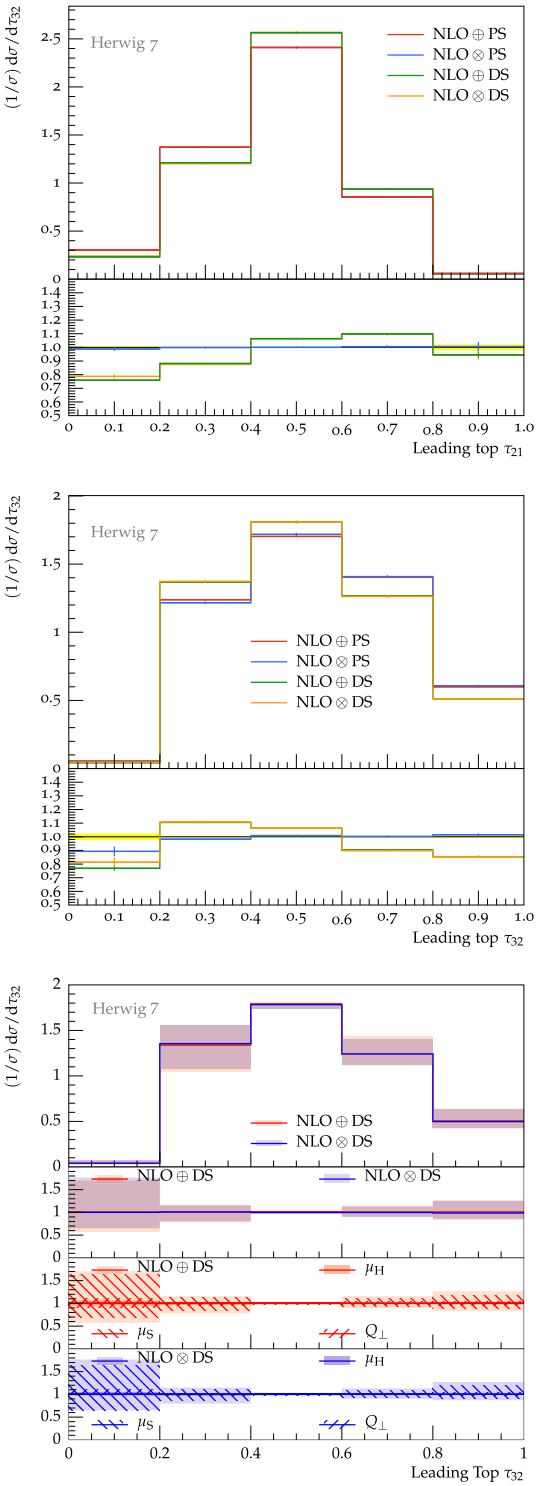


Fig. 16 The N-subjettiness ratios τ_{21} (top) and τ_{32} (middle and bottom) for the large-radius jet associated with the highest momentum top-quark in boosted $t\bar{t}$ events. The top and middle plots show comparisons of the angular-ordered (PS) and dipole (DS) showers with both the MC@NLO-type (NLO⊕) and Powheg-type (NLO⊗) matching schemes. The bottom plot shows the effects of scale variations on τ_{32} using the dipole shower with each of the NLO matching schemes

make the largest contribution to the uncertainty envelope, whereas the contributions from the other scale variations are negligible as we expect for an observable mostly sensitive to the top-quark decay. The choice of matching scheme is also found to have very little impact, except for the lowest τ_{32} bin. On the other hand, comparing the dipole shower and angular-ordered shower algorithms shows more significant differences, comparable to the uncertainty envelopes produced by scale variations.

In Fig. 17 we show predictions of the separation of the two hardest b-tagged jets in semi-leptonic $pp \to t\bar{t}$ events at a centre-of-collision energy of 8 TeV. The separation is defined as $\Delta R(j_{b1}, j_{b2}) = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, where $\Delta\phi$ and $\Delta\eta$ denote the difference in the azimuthal angle and pseudorapidity respectively of the hardest and second-hardest bottom-tagged jets. This observable is sensitive to both the simulation of the decay and to the direction of the top quarks that decay to produce the bottom quarks. We measure this distribution using a purpose-built analysis in which we require at least one final-state dressed lepton, electron or muon, with $p_\perp > 30$GeV and $|\eta| < 4.2$. Dressed leptons are created by clustering each bare lepton with any photons within a cone of $\Delta R = 0.1$ around the lepton. We also require at least two light-flavour jets and two bottom-tagged jets with $p_\perp > 30$GeV and $|\eta| < 4.2$. Additionally we implement a minimum missing transverse energy cut of 30 GeV, where the transverse energy of each visible outgoing particle is defined as $E_\perp = E\sin(\theta)$ where E and θ denote the energy and polar angle of the particle respectively, measured in the lab frame.

The dominant source of uncertainty on the LO predictions in the region $\Delta R < \pi$ is the variation of $Q_\perp$. This is because the relative orientation of the top quarks, and hence the separation of the bottom-tagged jets, is sensitive to hard radiation from the production process. The uncertainty envelopes on the NLO-matched predictions are in general smaller than those on the LO predictions, and there is no single dominant source of uncertainty. This is because the the hardest jet from the production process is simulated to LO, rather than parton-shower, accuracy. As we do not use the benchmark settings for the full-process simulation, which is intended for comparison to data, any conclusions about the effects of different shower or matching algorithms on these plots must be made with care. The uncertainty on this prediction due to scale variations is small and our findings suggest that most of the uncertainty is due to the sensitivity to the production process.

With relatively few experimental analyses that measure decay-process sensitive observables currently available, the

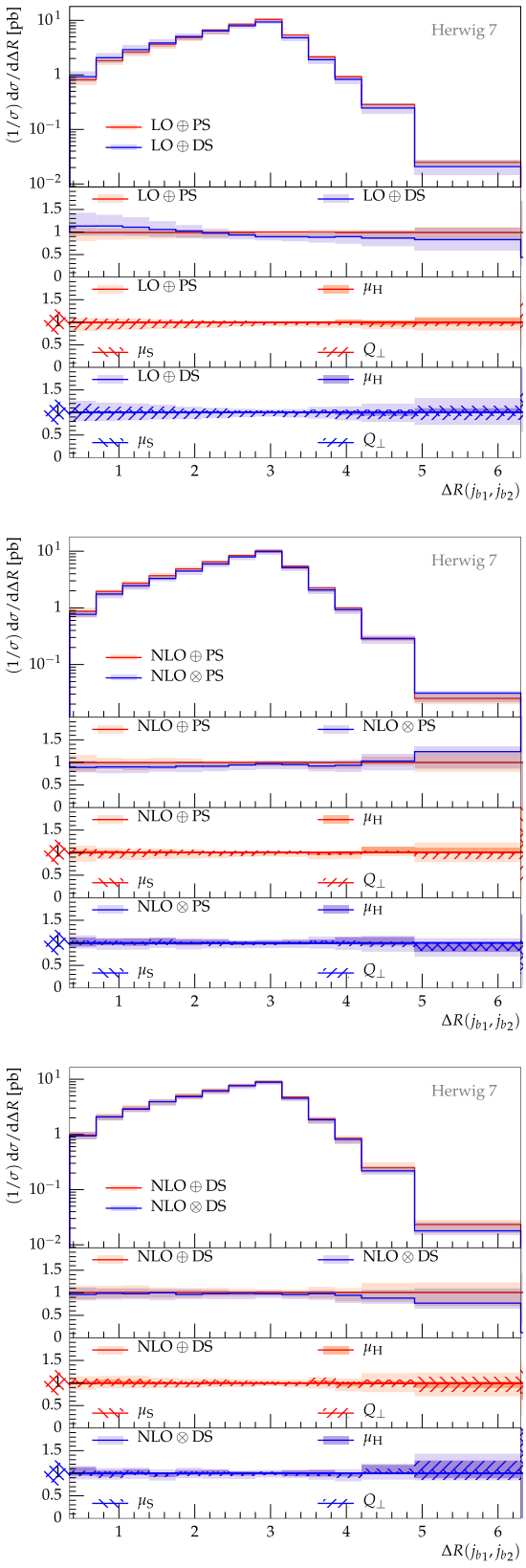


Fig. 17 The $\Delta R(b_1, b_2) = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ distribution, described in the text, simulated for semileptonic 8 TeV $pp \to t\bar{t}$ events

evaluation of the uncertainty on predictions of such observables is an area for future investigation.[16]

8 Summary and outlook

In this work we have presented a detailed study of NLO plus parton shower matched predictions for top pair production at the LHC in the HERWIG 7 event generator. We have considered various sources of uncertainty, including the matching algorithms themselves for which two options, a subtractive (MC@NLO-type) and multiplicative (Powheg-type) paradigm can be used within HERWIG 7, as well as all scale choices involved. We have not only considered NLO corrections to the production process, but also in the decay process. Both shower modules in HERWIG 7 are now able to handle radiation in both the production and the decay of top quarks, and we have used this paper as an opportunity to present a new treatment for radiation from heavy quarks in the dipole shower.

We have found that no single scale variation encompasses the entire set of independent variations, therefore all sources need to be considered to obtain a reliable estimate of the uncertainty on predictions. We have explicitly shown that NLO matching provides improvements over a LO plus parton shower simulation where expected. Higher jet multiplicities, however, do suffer from large uncertainties, even using NLO matching, a fact which should be considered when using tuned predictions. We have further considered boosted topologies, looking at observables that are sensitive to the decay process, focusing on N-subjettiness ratios which highlight the internal structure of the jets.

Particular attention has been paid to the choice of the hard veto scale and the profile scale. This is an ambiguity in matching algorithms which has not been addressed extensively in the literature (with some investigations presented in Ref. [11], and a more elaborate leading-order analysis presented in Ref. [24]) but plays an important role in the handling of real-emission corrections present in the NLO matching. Inappropriate choices can lead to artificially suppressed or enhanced radiation, and we have found that scales which identify the hard objects in the process provide the most reliable results.

The main purpose of this work is to highlight the uncertainties and ambiguities associated with NLO matching, which need to be compared between different shower and

[16] In Sect. 4.2.4 we use the process $e^+e^- \to t\bar{t}$ at production threshold to validate the new decay kernels in the dipole shower, where the reduced radiation in the production process allows us to construct analyses that are very sensitive to the treatment of the top quark decay (see Fig. 5). For $pp \to t\bar{t}$ we found the ability to assess the impact of the matrix element correction to the first emission from top quark decays is limited by the increased sensitivity to the production process which dominates most observables.

matching algorithms. The HERWIG 7 event generator provides unique capabilities to quantify the differences between predictions obtained using different setups and to benchmark variations against each other. These sources of uncertainty should be taken into account when comparing predictions against data, also in view of an improved simulation based on multi-jet merging, which can more reliably predict higher jet multiplicities. Being beyond the scope of this paper, a multi-jet merged study, considering a similar set of variations, would thus be a logical follow-up for future work.

Acknowledgements SW acknowledges support from a STFC studentship. KC and SP are grateful to The University of Manchester for kind hospitality during part of this work. This work was supported in part by the European Union as part of the FP7 Marie Curie Initial Training Network MCnetITN (PITN-GA-2012-315877). In addition this work has received funding from the European Union's Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (Grant agreement no. 722104). SP acknowledges partial support by the COST Action CA16201 PARTICLEFACE, and by a FP7 Marie Curie Intra European Fellowship under Grant Agreement PIEF- GA-2013-628739, when part of this work was started. CR acknowledges current support by the European Union's Horizon 2020 research and innovative programme, under Grant agreement No. 668679, and recent support by the U.S. Department of Energy, under the Grant agreements DE-SC0010102 and DE-FG02-13ER41942 when part of this work was done, as well as partial support by the German Federal Ministry of Education and Research (BMBF) when part of this work was started. We would like to thank the other members of the HERWIG collaboration for their continuous support. In particular we would like to thank Johannes Bellm and Stefan Gieseke for valuable discussions, as well as Martin Stoll for his initial contribution to massive dipoles.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The simulated data exists in the form of Rivet histograms, in a private repository, but can be made available on request if needed. The experimental data is publicly available as Rivet analyses, associated to the correspondingly cited experimental analyses.]

Funded by SCOAP[3].

A Dipole shower kinematics

A.1 Alternative formulation for the final–final dipole kinematics

The physical momenta of the partons following a splitting from a massive final–final dipole, written in terms of the splitting variables z_i and $y_{ij,k}$, are

$$q_i = A_i\, Q + k_\perp + B_i\, v_\parallel, \tag{67a}$$

$$q_j = A_j\, Q - k_\perp + B_j\, v_\parallel, \tag{67b}$$

$$q_k = A_k\, Q + B_k\, v_\parallel, \tag{67c}$$

where

$$A_i = \frac{1}{s}\left[m_i^2 + \frac{\bar{s}}{2}\left(y_{ij,k} + z_i(1-y_{ij,k})\right)\right], \tag{68a}$$

$$A_j = \frac{1}{s}\left[m_j^2 + \frac{\bar{s}}{2}\left(1 - z_i(1-y_{ij,k})\right)\right], \tag{68b}$$

$$A_k = \frac{1}{s}\left[m_k^2 + \frac{\bar{s}}{2}(1-y_{ij,k})\right], \tag{68c}$$

$$B_i = \frac{1}{B_k}\left(s A_i A_k - \frac{\bar{s}}{2} z_i (1-y_{ij,k})\right), \tag{68d}$$

$$B_j = \frac{1}{B_k}\left(s A_j A_k - \frac{\bar{s}}{2}(1-z_i)(1-y_{ij,k})\right), \tag{68e}$$

$$B_k = \sqrt{\frac{1}{s}\left(m_k^2 + \frac{\bar{s}}{2}(1-y_{ij,k})\right)^2 - m_k^2}, \tag{68f}$$

and the 4-vector

$$v_\parallel = \sqrt{\frac{4s}{\lambda\left(s, m_k^2, m_{ij}^2\right)}}\left(\tilde{p}_k - \frac{Q\cdot\tilde{p}_k}{s} Q\right), \tag{69}$$

is expressed using the Kallen function

$$\lambda(x,y,z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz\,. \tag{70}$$

Note that while it is trivial to write an expression $p_\perp = p_\perp(z_i, y_{ij,k})$, this expression is cubic in $y_{ij,k}$ which leads to a complicated analytic equation for $y_{ij,k} = y_{ij,k}(p_\perp, z_i)$.

A.2 Final–final dipole kinematics

A.2.1 Completing the formulation

In order to complete our formulation of the splitting kinematics in Sect. 3.3.1 we require expressions for the scaling parameters x_k and x_{ij} in terms of the variables $p_\perp$ and z. We first write an expression for the virtuality of the emitter and emission produced in the splitting

$$Q_{ij}^2 = \left(q_i + q_j\right)^2 = \frac{1}{z(1-z)}\left[p_\perp^2 + (1-z)m_i^2 + z m_j^2\right]. \tag{71}$$

Defining the variables λ_{ij} and λ_k as

$$\lambda_k = 1 + \frac{m_k^2}{s_{ij,k}}, \quad \lambda_{ij} = 1 + \frac{m_{ij}^2}{s_{ij,k}}\,, \tag{72}$$

we derive the following expressions for the scaling parameters

$$x_{ij} = 1 - \frac{m_k^2}{s_{ij,k}}\frac{(1-x_k)}{x_k}\,, \tag{73}$$

$$x_k = \frac{1}{2\lambda_k}\left[\left(\lambda_{ij}\lambda_k + \frac{m_k^2}{s_{ij,k}} - \frac{Q_{ij}^2}{s_{ij,k}}\right) \pm \sqrt{\left(\lambda_{ij}\lambda_k + \frac{m_k^2}{s_{ij,k}} - \frac{Q_{ij}^2}{s_{ij,k}}\right)^2 - 4\lambda_{ij}\lambda_k\frac{m_k^2}{s_{ij,k}}}\right]. \tag{74}$$

Finally we require expressions for the splitting variables z_i and $y_{ij,k}$ in terms of the variables $p_\perp$ and z. We write $y_{ij,k}$ as

$$y_{ij,k} = \frac{1}{\bar{s}z(1-z)}\left[p_\perp^2 + (1-z)^2 m_i^2 + z^2 m_j^2\right], \tag{75}$$

where the invariant quantity $\bar{s}$ is

$$\bar{s} = s - m_i^2 - m_j^2 - m_k^2 . \tag{76}$$

In order to express z_i in terms of $p_\perp$ and z we write

$$z_i = \frac{2q_i \cdot q_k}{(1-y_{ij,k})\bar{s}}, \tag{77}$$

where the denominator can be written as

$$(1-y_{ij,k})\bar{s} = \frac{1}{z(1-z)}\left[z(1-z) - (1-z)^2 m_i^2 - z^2 m_j^2 - p_\perp^2\right], \tag{78}$$

and the numerator is given by the expression

$$2q_i \cdot q_k = z x_{ij} x_k s_{ij,k} + \frac{m_k^2}{z x_{ij} x_k s_{ij,k}}\left(p_\perp^2 + m_i^2\right). \tag{79}$$

A.2.2 Phase-space limits

In order to allow us to efficiently generate values for $p_\perp$ and z according to the splitting kernels we need to express the single-particle emission phase space and the limits on it in terms of these variables.

The limits on the dipole splitting variables z_i and $y_{ij,k}$ are given in Ref. [34] and we include them here to provide a complete reference,

$$y_{ij,k,-} = \frac{2m_i m_j}{\bar{s}}, \tag{80}$$

$$y_{ij,k,+} = 1 - \frac{2m_k(\sqrt{s}-m_k)}{\bar{s}}, \tag{81}$$

$$z_{i,\pm}(y_{ij,k}) = \frac{2m_i^2 + \bar{s}y_{ij,k}}{2\left[m_i^2 + m_j^2 + \bar{s}y_{ij,k}\right]}(1 \pm v_{ij,i}v_{ij,k}), \tag{82}$$

where the relative velocities $v_{ij,k}$ and $v_{ij,i}$ are expressed as functions of $y_{ij,k}$,

$$v_{ij,k} = \frac{\sqrt{\left[2m_k^2 + \bar{s}(1-y_{ij,k})\right]^2 - 4m_k^2 s}}{\bar{s}(1-y_{ij,k})}, \tag{83}$$

$$v_{ij,i} = \frac{\sqrt{\bar{s}^2 y_{ij,k}^2 - 4m_i^2 m_j^2}}{\bar{s}y_{ij,k} + 2m_i^2}. \tag{84}$$

The limits on $p_\perp$ and z follow from the inequality $y_{ij,k} < y_{ij,k,+}$,

$$p_{\perp,\max} = \frac{1}{2(\sqrt{s}-m_k)}\sqrt{\lambda\left(m_i^2, m_j^2, (\sqrt{s}-m_k)^2\right)}, \tag{85}$$

$$z_\pm = \frac{1}{2\left(\sqrt{s}-m_k\right)^2}\left[m_i^2 - m_j^2 + (\sqrt{s}-m_k)^2\right. \tag{86}$$

$$\left. \pm\sqrt{\lambda\left(m_i^2, m_j^2, (\sqrt{s}-m_k)^2\right)}\sqrt{1 - \frac{p_\perp^2}{p_{\perp,\max}^2}}\right]. \tag{87}$$

A.2.3 The single-particle emission phase space

The single-particle emission phase space required to express the branching probability in Eq. (18) is written as

$$\mathrm{d}q_j = \frac{1}{16\pi^2}\frac{\bar{s}^2}{\sqrt{\lambda\left(s, m_{ij}^2, m_k^2\right)}}\left(1-y_{ij,k}\right)\mathrm{d}y_{ij,k}\mathrm{d}z_i\frac{\mathrm{d}\phi}{2\pi}. \tag{88}$$

As we consider only spin-averaged kernels the azimuthal angle, ϕ, is averaged over in the phase-space integration and we do not consider it explicitly in the following discussion.

Using the above expression for the single-particle phase space we express the branching probability as

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{16\pi^2}\langle V_{ij,k}\left(z_i, y_{ij,k}\right)\rangle\frac{1}{\left(1 + \frac{m_i^2+m_j^2-m_{ij}^2}{\bar{s}y_{ij,k}}\right)} \times \frac{\bar{s}}{\sqrt{\lambda\left(s, m_{ij}^2, m_k^2\right)}}\left(1-y_{ij,k}\right)\frac{\mathrm{d}y_{ij,k}}{y_{ij,k}}\mathrm{d}z_i, \tag{89}$$

where we can express the phase-space integration in terms of the generated variables using the replacement

$$\frac{\mathrm{d}y_{ij,k}}{y_{ij,k}}\mathrm{d}z_i \to \left[\frac{p_\perp^2}{p_\perp^2 + (1-z)^2 m_i^2 + z^2 m_j^2}\right] \times \left[1 - 2\frac{1}{\bar{s}(1-y_{ij,k})}\frac{m_k^2 Q_{ij}^2}{x_{ij}x_k s_{ij,k}}\right]\frac{\mathrm{d}p_\perp^2}{p_\perp^2}\mathrm{d}z . \tag{90}$$

A.3 Final–initial dipole kinematics

A.3.1 Phase-space limits

The upper limit on $x_{ij,b}$ is,

$$x_{ij,b,+} = \frac{s_{ij,b}}{s_{ij,b} - m_{ij}^2 + (m_i + m_j)^2} \,. \tag{91}$$

We can derive a lower limit on $x_{ij,b}$. We first write the momentum of the incoming proton as P and the proton momentum-fraction carried by the spectator prior to the splitting as x_s. We can write

$$q_b = \frac{1}{x_{ij,b}} \tilde{p}_b = \frac{1}{x_{ij,b}} (x_s P) < P \,, \tag{92}$$

hence we require

$$x_{ij,b} > x_s \,. \tag{93}$$

From the inequality in Eq. (93) we derive the following limits on the variables $p_\perp$ and z

$$p^2_{\perp,\max} = \frac{s'_{ij,b}}{4} \lambda \left(1, \frac{m_i^2}{s'_{ij,b}}, \frac{m_j^2}{s'_{ij,b}}\right) \,, \tag{94}$$

$$z_\pm = \frac{1}{2} \left[1 + \frac{m_i^2 - m_j^2}{s'_{ij,b}} \pm \sqrt{\lambda \left(1, \frac{m_i^2}{s'_{ij,b}}, \frac{m_j^2}{s'_{ij,b}}\right)} \sqrt{1 - \frac{p_\perp^2}{p^2_{\perp,\max}}} \right] \,, \tag{95}$$

where λ is the standard Kallen function and for convenience we have defined the modified invariant

$$s'_{ij,b} = s_{ij,b} \left(\frac{1 - x_s}{x_s} \right) + m_{ij}^2 \,. \tag{96}$$

A.3.2 The single-particle emission phase space

The single-particle emission phase space required to express the branching probability in Eq. (28) is written as

$$\mathrm{d}q_j = \frac{1}{16\pi^2} 2\tilde{p}_{ij} \cdot q_b \mathrm{d}z_i \mathrm{d}x_{ij,b} \frac{\mathrm{d}\phi}{2\pi} \,. \tag{97}$$

As we consider only spin-averaged kernels the azimuthal angle is averaged over in the phase space integration.

Using the above expression for the single-particle phase space we express the branching probability as

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{16\pi^2} \frac{f_b(x_s/x_{ij,b})}{f_b(x_s)} \langle V_{ij}^b \left(z_i, x_{ij,b}\right) \rangle \times \frac{1}{x_{ij,b}(1 - x_{ij,b})} \mathrm{d}z_i \mathrm{d}x_{ij,b}. \tag{98}$$

Noting that $z_i = z$ we can express the phase-space integration in terms of the generated variables using the replacement

$$l \quad \frac{1}{x_{ij,b}(1 - x_{ij,b})} \mathrm{d}z_i \mathrm{d}x_{ij,b} \to \left[\frac{p_\perp^2}{p_\perp^2 + (1-z)m_i^2 + z m_j^2 - z(1-z)m_{ij}^2} \right] \frac{\mathrm{d}p_\perp^2}{p_\perp^2} \mathrm{d}z. \tag{99}$$

A.4 Initial–final dipole kinematics

A.4.1 Phase-space limits

The lower and upper limits on u_j are

$$u_{j,+} = 0 \tag{100}$$

$$u_{j,+} = \frac{1 - x_{jk,a}}{1 - x_{jk,a}(1 - m_k^2/s_{aj,k})} \,, \tag{101}$$

and the upper limit on $x_{jk,a}$ is

$$x_{jk,a,+} = 1. \tag{102}$$

Following an analogous argument to that used to derive the inequality in Eq. (93) we derive a lower limit for $x_{jk,a}$

$$x_{jk,a} > x_e \,, \tag{103}$$

where x_e is the proton momentum-fraction carried by the emitter prior to the splitting.

From the inequality in Eq. (103) we derive the following limits on the variables $p_\perp$ and z

$$p^2_{\perp,\max} = \frac{{s'_{aj,k}}^2}{4} \left[\frac{1}{m_k^2 + s'_{aj,k}} \right] \,, \tag{104}$$

$$z_\pm = \frac{1}{2} \left[(1 + x_e) \pm (1 - x_e) \sqrt{1 - \frac{p_\perp^2}{p^2_{\perp,\max}}} \right] \,, \tag{105}$$

where for convenience we have defined the rescaled invariant

$$s'_{aj,k} = s_{aj,k} \left(\frac{1 - x_e}{x_e} \right) \,. \tag{106}$$

A.4.2 The single-particle emission phase space

The single-particle emission phase space required to express the branching probability in Eq. (36) is written as

$$\mathrm{d}q_j = \frac{1}{16\pi^2} 2q_a \cdot \tilde{p}_k \mathrm{d}u_j \mathrm{d}x_{jk,a} \frac{\mathrm{d}\phi}{2\pi} \,. \tag{107}$$

As we consider only spin-averaged kernels the azimuthal angle is averaged over in the phase-space integration.

Using the above expression for the single-particle phase space we express the branching probability as

$$\mathrm{d}\mathcal{P}_{\text{branching}} = \frac{1}{16\pi^2}\frac{f_a(x_e/x_{jk,a})}{\tilde{f}_{aj}(x_e)}\langle V_k^{aj}\left(u_j, x_{jk,a}\right)\rangle \times \frac{1}{u_j}\frac{1}{x_{jk,a}}\mathrm{d}u_j\mathrm{d}x_{jk,a}\,, \tag{108}$$

where we can express the phase-space integration in terms of the generated variables using the replacement

$$\frac{1}{u_j}\frac{1}{x_{jk,a}}\mathrm{d}u_j\mathrm{d}x_{jk,a} \to \left[u_j + x_{jk,a} - u_j x_{jk,a}\left(1 - \frac{m_k^2}{s_{aj,k}}\right)\right]^{-1}\frac{\mathrm{d}p_\perp^2}{p_\perp^2}\mathrm{d}z\,. \tag{109}$$

B Generation cut for boosted top analyses

A difficulty for high momentum top quark simulation is that of CPU time required to populate the high-$p_\perp$ region of phase space targeted by these analyses. Given the steeply falling $t\bar{t}$ cross-section as a function of the transverse momentum of the top quark, analyses targeting boosted topologies are typically targeting $\sim$ 1% of the total phase space or less. Simulating the inclusive phase-space can therefore be very inefficient, requiring orders of magnitude more events to be fully simulated than are actually of interest. In order to reduce the resources required for simulation a simple cut mechanism is available in the MATCHBOX framework of HERWIG 7. This mechanism makes several options available to improve the efficiency of producing high-$p_\perp$ top quarks.

We can use the existing cut infrastructure in HERWIG 7 to implement a generation cut to enhance the production of events that include a boosted top quark or antiquark. We create a 'MatchboxFactoryMatcher' that identifies top quarks and antiquarks and associate this with the existing 'JetFinder' and 'JetCuts' objets. Then we assign the existing 'FirstJet' and 'SecondJet' objects to the 'JetRegions' of the 'JetCuts' object. As we have assigned the matcher of the 'JetCuts' object to be our new top quark matcher, the 'FirstJet' and 'SecondJet' actually identify top quarks. Therefore we can set transverse momentum and rapidity cuts on the top quarks by setting the cuts on the 'FirstJet' and 'SecondJet'. The code snippet required to do this is included below.

Example runs with cuts on the transverse momentum of the leading top quark of 200, 300 and 600 GeV were performed to test the efficiency. They showed no significant change in the distributions of weights. For a centre-of-mass energy of 13 TeV kinematic bins well beyond the cut values increased their overall statistics by factors of $\sim$ 5, $\sim$ 20 and $\sim$ 500 respectively for the same total number of events, and relative errors were reduced accordingly. No appreciable impact on the computing time per event was found, allowing a significant reduction in computing power to achieve the same or better statistical power.

A code snippet of an input card to produce similar cuts is given below.

```
##################################################
## Cut for boosted top analyses
##################################################
cd /Herwig/MatrixElements/Matchbox

# Create a new particle group consisting of top
# quarks and antiquarks
do Factory:StartParticleGroup ttbar
insert Factory:ParticleGroup 0 /Herwig/
Particles/t
insert Factory:ParticleGroup 0 /Herwig/
Particles/tbar
do Factory:EndParticleGroup

# Create a new matcher and associate it with the
# top quark particle group
create Herwig::MatchboxFactoryMatcher
TopAntiTopMatcher
set TopAntiTopMatcher:Factory /Herwig/
MatrixElements/Matchbox/Factory
set TopAntiTopMatcher:Group ttbar

# Set the matcher of the JetFinder and JetCuts
# to the new top quark matcher
set /Herwig/Cuts/JetFinder:UnresolvedMatcher
TopAntiTopMatcher
set /Herwig/Cuts/JetCuts:UnresolvedMatcher
TopAntiTopMatcher

cd /Herwig/Cuts

# This snippet sets up JetFinder and JetCuts
read Matchbox/DefaultPPJets.in

# Set up the JetRegions and cuts
# Note: FirstJet and SecondJet are actually top
# quarks/antiquarks
insert JetCuts:JetRegions 0 FirstJet
insert JetCuts:JetRegions 1 SecondJet

set FirstJet:PtMin 0.*GeV
do FirstJet:YRange -5.0 5.0

set SecondJet:PtMin 0.*GeV
do SecondJet:YRange -5.0 5.0
```

References

1. M. Czakon, P. Fiedler, A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013). arXiv:1303.6254
2. CMS, V. Khachatryan et al., Phys. Rev. **D93**, 072004 (2016). arXiv:1509.04044

3. ATLAS, M. Aaboud et al., Phys. Lett. **B761**, 350 (2016). arXiv:1606.02179
4. ATLAS, G. Aad et al., JHEP **06**, 100 (2015). arXiv:1502.05923
5. T. Gleisberg et al., JHEP **02**, 007 (2009). arXiv:0811.4622
6. M. Bähr et al., Eur. Phys. J. C **58**, 639 (2008). arXiv:0803.0883
7. T. Sjöstrand et al., Comput. Phys. Commun. **191**, 159 (2015). arXiv:1410.3012
8. S. Frixione, B.R. Webber, JHEP **06**, 029 (2002). arXiv:hep-ph/0204244
9. S. Frixione, P. Nason, C. Oleari, JHEP **11**, 070 (2007). arXiv:0709.2092
10. S. Plätzer, S. Gieseke, Eur. Phys. J. C **72**, 2187 (2012). arXiv:1109.6256
11. S. Hoeche, F. Krauss, M. Schonherr, F. Siegert, JHEP **1209**, 049 (2012). arXiv:1111.1220
12. S. Plätzer (2012). arXiv:1211.5467
13. L. Lonnblad, S. Prestel (2012). arXiv:1211.4827
14. L. Lonnblad, S. Prestel (2012). arXiv:1211.7278
15. J. Bellm, S. Gieseke, S. Plätzer, Eur. Phys. J. C **78**, 244 (2018). arXiv:1705.06700
16. K. Hamilton, P. Nason, G. Zanderighi, JHEP **1210**, 155 (2012). arXiv:1206.3572
17. T. Jezo, J.M. Lindert, P. Nason, C. Oleari, S. Pozzorini, Eur. Phys. J. C **76**, 691 (2016). arXiv:1607.04538
18. S. Fleming, A.H. Hoang, S. Mantry, I.W. Stewart, Phys. Rev. D **77**, 074010 (2008). arXiv:hep-ph/0703207
19. M. Beneke, P. Marquard, P. Nason, M. Steinhauser, Phys. Lett. B **775**, 63 (2017). arXiv:1605.03609
20. P. Nason, The Top Mass in Hadronic Collisions (2017). arXiv:1712.02796
21. A. H. Hoang, S. Plätzer, D. Samitz (2018). arXiv:1807.06617
22. G. Heinrich et al., JHEP **07**, 129 (2018). arXiv:1709.08615
23. G. Corcella, R. Franceschini, D. Kim, Nucl. Phys. B **929**, 485 (2018). arXiv:1712.05801
24. J. Bellm, G. Nail, S. Plätzer, P. Schichtel, A. Siódmok, Eur. Phys. J. C **76**, 665 (2016). arXiv:1605.01338
25. J. Alwall et al., JHEP **07**, 079 (2014). arXiv:1405.0301
26. F. Cascioli, P. Maierhofer, S. Pozzorini, Phys. Rev. Lett. **108**, 111601 (2012). arXiv:1111.5206
27. http://mcfm.fnal.gov
28. J.M. Campbell, R.K. Ellis, Phys. Rev. D **60**, 113006 (1999). arXiv:hep-ph/9905386
29. J.M. Campbell, R.K. Ellis, C. Williams, JHEP **07**, 018 (2011). arXiv:1105.0020
30. D. Rauch, Automated NLO calculations with Massive Quarks in Herwig++, Master's thesis, KIT (2014)
31. P. Nason, JHEP **11**, 040 (2004). hep-ph/0409146
32. S. Gieseke, P. Stephens, B.R. Webber, JHEP **12**, 045 (2003). arXiv:hep-ph/0310083
33. S. Catani, M.H. Seymour, Nucl. Phys. B **485**, 291 (1997). arXiv:hep-ph/9605323
34. S. Catani, S. Dittmaier, M.H. Seymour, Z. Trocsanyi, Nucl. Phys. B **627**, 189 (2002). arXiv:hep-ph/0201036
35. S. Plätzer, S. Gieseke, JHEP **01**, 024 (2011). arXiv:0909.5593
36. G. Marchesini, B.R. Webber, Nucl. Phys. B **330**, 261 (1990)
37. G. Corcella et al., JHEP **01**, 010 (2001). arXiv:hep-ph/0011363
38. M. Stoll, Diploma Thesis, Universität Karlsruhe (2012)
39. S. Schumann, F. Krauss, JHEP **03**, 038 (2008) **(0709.1027)**
40. S. Catani, S. Dittmaier, Z. Trocsanyi, Phys. Lett. B **500**, 149 (2001). arXiv:hep-ph/0011222
41. S. Hoeche, S. Schumann, F. Siegert, Phys. Rev. D **81**, 034026 (2010). arXiv:0912.3501
42. P. Richardson, A. Wilcock, Eur. Phys. J. C **74**, 2713 (2014). arXiv:1303.4563
43. K. Hamilton, P. Richardson, JHEP **07**, 010 (2006). arXiv:hep-ph/0603034
44. G. Corcella, M.H. Seymour, Phys. Lett. B **442**, 417 (1998). arXiv:hep-ph/9809451
45. K. Hamilton, P. Richardson, JHEP **02**, 069 (2007). arXiv:hep-ph/0612236
46. K. Melnikov, M. Schulze, A. Scharf, Phys. Rev. D **83**, 074013 (2011). arXiv:1102.1967
47. K. Melnikov, A. Scharf, M. Schulze, Phys. Rev. D **85**, 054002 (2012). arXiv:1111.4991
48. L. Basso, S. Dittmaier, A. Huss, L. Oggero, Eur. Phys. J. C **76**, 56 (2016). arXiv:1507.04676
49. J. Bellm, Eur. Phys. J. C **78**, 601 (2018). arXiv:1801.06113
50. S. Plätzer, M. Sjodahl, J. Thorén (2018). arXiv:1808.00332
51. L.H. Orr, T. Stelzer, W.J. Stirling, Phys. Rev. D **56**, 446 (1997). arXiv:hep-ph/9609246
52. S. Catani, Y.L. Dokshitzer, M. Olsson, G. Turnock, B.R. Webber, Phys. Lett. B **269**, 432 (1991)
53. M. Cacciari, G.P. Salam, G. Soyez, Eur. Phys. J. C **72**, 1896 (2012). arXiv:1111.6097
54. P. Richardson, S. Webster (2018). arXiv:1807.01955
55. S. Plätzer, M. Sjodahl, JHEP **07**, 042 (2012). arXiv:1201.0260
56. S. Plätzer, Eur. Phys. J. C **72**, 1929 (2012). arXiv:1108.6182
57. S. Alioli, P. Nason, C. Oleari, E. Re, JHEP **1006**, 043 (2010). arXiv:1002.2581
58. G. Heinrich, S.P. Jones, M. Kerner, G. Luisoni, E. Vryonidou, JHEP **08**, 088 (2017). arXiv:1703.09252
59. M. Czakon, D. Heymes, A. Mitov, JHEP **04**, 071 (2017). arXiv:1606.03350
60. M. Czakon, A. Mitov, Comput. Phys. Commun. **185**, 2930 (2014). arXiv:1112.5675
61. A. Buckley et al., Comput. Phys. Commun. **184**, 2803 (2013). arXiv:1003.0694
62. M. Cacciari, G.P. Salam, G. Soyez, JHEP **04**, 063 (2008). arXiv:0802.1189
63. ATLAS, G. Aad et al., JHEP **01**, 020 (2015). arXiv:1407.0891
64. ATLAS, G. Aad et al., Eur. Phys. J. **C76**, 538 (2016). arXiv:1511.04716
65. CMS, V. Khachatryan et al., Phys. Rev. **D94**, 052006 (2016). arXiv:1607.00837
66. ATLAS, G. Aad et al., Eur. Phys. J. **C72**, 2043 (2012). arXiv:1203.5015
67. J. Thaler, K. Van Tilburg (2011). arXiv:1011.2268