Eur. Phys. J. C (2024) 84:212
https://doi.org/10.1140/epjc/s10052-024-12562-7

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

On heavy-flavour jets with Soft Drop

Simone Caletti[1,a], **Andrea Ghira**[2,b], **Simone Marzani**[2,c]

[1] Institute for Theoretical Physics, ETH, 8093 Zurich, Switzerland
[2] Dipartimento di Fisica, Università di Genova and INFN, Sezione di Genova, Via Dodecaneso 33, 16146 Genoa, Italy

Received: 22 December 2023 / Accepted: 14 February 2024 / Published online: 29 February 2024

Abstract We study hadronic jets that are tagged as heavy-flavoured, i.e. they contain either beauty or charm. In particular, we consider heavy-flavour jets that have been groomed with the Soft Drop algorithm. In order to achieve a deeper understanding of these objects, we apply resummed perturbation theory to jets initiated by a massive quark and we perform analytic calculations for two variables that characterise Soft Drop jets, namely the opening angle and the momentum fraction of the splitting that passes Soft Drop. We compare our findings to Monte Carlo simulations. Furthermore, we investigate the correlation between the Soft Drop energy fraction and alternative observables that aim to probe heavy-quark fragmentation functions.

1 Introduction and motivation

Beauty (b) and charm (c) quarks are often considered "heavy flavours" because their masses are above the proton mass, $m_b = 4.2$ GeV and $m_c = 1.3$ GeV, respectively. However, their mass is not so large, compared to the typical scale of hadron formation, $\Lambda \simeq 1$ GeV, so hadronisation occurs. This does not happen for the top quark, which has a mass 170 times bigger than the proton mass. This value is so large that its lifetime is smaller than the hadronisation scale.

Heavy flavours constitute a window on two mechanisms that provide ordinary matter with mass: electroweak symmetry breaking and the binding energy of strong interactions. On the one hand, they play a crucial role in studies of the Higgs boson and, on the other hand, they constitute a bridge between perturbative and non-perturbative Quantum Chromo Dynamics (QCD). For these reasons, they have been (and still are) the subject of detailed theoretical and experimental studies.

In this work, we focus on processes in which heavy quarks are produced. From a theoretical point of view, calculations for identified heavy flavours can be performed essentially because the quark mass sets a perturbative scale for the running coupling and, at the same time, removes collinear singularities. From an experimental viewpoint, the lifetime of B (or D) hadrons is long enough so that their decay happens away from the interaction point. Dedicated b- and c-tagging techniques that exploit this property to identify B and D hadrons, or b and c jets, are widely used in collider experiments.

Flavour physics has been studied for decades, both in the quark and lepton sectors. However, the recent development of Infra-Red and Collinear (IRC) safe flavour-jet algorithms [1–5] opens up the possibility to set up a yet-unexplored flavour physics program that exploits jets and their substructure at the Large Hadron Collider (LHC).

Resummed calculations for jets initiated by heavy quarks have been first performed in the context of studies focussing on B decays [6–9] and top jets [10–15], exploiting effective field theories. However, to the best of our knowledge, there exists only a handful of studies that exploit modern jet substructure techniques to study heavy flavours [16–23]. The main goal of these studies is to investigate the so-called dead-cone effect [24,25], i.e. the fact that colour radiation around heavy quarks is suppressed. Remarkably, this effect has been recently measured by the ALICE collaboration at the LHC [26].

In contrast, there exists a rather extensive literature dedicated to studying the properties of a reconstructed B (or D) hadron, such as its energy or its transverse momentum. The theoretical description of these observables is usually based on heavy-quark fragmentation functions, which can be computed in perturbative QCD. State-of-the-art predictions include the resummation of different classes of logarithmic corrections, e.g. mass logarithms and end-point log-

[a] e-mail: scaletti@phys.ethz.ch (corresponding author)
[b] e-mail: andrea.ghira@ge.infn.it
[c] e-mail: simone.marzani@ge.infn.it

arithms, see e.g. [27–34]. In this context, two of us have recently developed a theoretical framework to consistently resum both mass and soft logarithms [35,36]. In this study, we will heavily rely on these results and we will apply them to jet-based observables. This way, we will be able to obtain theoretical predictions that resum both the logarithms of the observable we want to study and the logarithms of the ratio of the heavy-quark mass to the jet transverse momentum.

In this study, we investigate the possibility of exploiting a widely used jet substructure technique, namely Soft Drop [37], to study heavy-flavour jets. The Soft Drop algorithm has been extensively studied and Soft Drop jets are routinely used in experimental analyses, both in the context of measurements and searches for new physics. In particular, we consider heavy-flavour-tagged jets and focus on the angular separation (θ_g) and momentum fraction (z_g) of the first splitting that passes Soft Drop. The former directly measures the angular resolution of the groomed jet and, therefore, it gives us direct access to the dead cone. The latter, instead, allows us to probe the heavy-quark splitting function [38,39].

These observables have been recently measured on c-jets by the ALICE collaboration [40] and our study constitutes the first step towards a first-principle description of the data. A direct comparison to the ALICE data, however, goes beyond the scope of this paper, as it would require matching the resummed expressions to fixed-order matrix elements, as well as accounting for the fact that ALICE reconstructs c-jets from exclusive D-meson decays.

The paper is organised as follows. We start by reviewing the Soft Drop algorithm in Sect. 2. In Sect. 3 we perform the next-to-leading logarithmic (NLL) resummation of the θ_g distribution, including heavy-quark mass effects and compare it Monte Carlo parton-shower simulations. In Sect. 4 we study the z_g distribution, both from first-principle and in Monte Carlo simulations and study its correlation to a standard momentum fraction variable used in the context of fragmentation functions. Finally, we draw our conclusions in Sect. 5. Details of the calculations are collected in the Appendices.

2 Soft Drop

Soft Drop [37] is a grooming algorithm that recursively removes soft-wide angle constituents from a jet. The Soft Drop procedure starts by re-clustering a given jet (typically an anti-k_t [41] jet) with radius R_0 and transverse momentum p_t with the Cambridge–Aachen (C/A) algorithm [42,43]. Soft Drop then parses the resulting angular-ordered branching history, grooming away the softer branch, until a branch that satisfies the condition

$$\frac{\min(p_{t1}, p_{t2})}{p_{t1}+p_{t2}} > z_c \left(\frac{\Delta_{12}}{R_0}\right)^{\beta}, \qquad \Delta_{12} = \sqrt{(y_1-y_2)^2+(\phi_1-\phi_2)^2}, \tag{1}$$

is found. In the expression above, 1 and 2 denote the branches at a given step in the clustering, p_{ti} are the corresponding transverse momenta, and Δ_{12} is their rapidity-azimuth separation. If the condition above is never satisfied, we can either remove the jet from consideration ("tagging mode") or leave it as the final Soft Drop jet ("grooming mode").

The kinematics of the first branch that satisfies (1) defines the groomed jet radius θ_g and the groomed momentum sharing z_g:

$$\theta_g = \frac{\Delta_{12}}{R_0}, \quad z_g = \frac{\min(p_{t1}, p_{t2})}{p_{t1}+p_{t2}}. \tag{2}$$

The θ_g distribution is IRC safe and it was first studied, for light jets, in [37] and then resummed to next-to-leading logarithmic (NLL) accuracy in [44]. The momentum sharing z_g is IRC safe for $\beta < 0$ but only Sudakov safe for $\beta \geq 0$ [38]. The NLL calculation of the z_g distribution was performed in [45].

3 The θ_g distribution

In this section, we describe the resummation of the θ_g distribution for a jet originated by a massive quark. We follow the general strategy described in Ref. [37] and we, therefore, consider the resummation of the cumulative distribution, normalised to the Born cross section,

$$\Sigma(\theta_g^2) = \frac{1}{\sigma_0}\int_0^{\theta_g^2} \frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g'^2}\,\mathrm{d}\theta_g'^2, \tag{3}$$

exploiting Lund diagrams [46].

3.1 Lund plane geography

Lund diagrams are a useful way to represent the available phase space for the emission of soft and collinear gluons off a hard dipole. The Lund plane is usually represented in terms of pairs of the logarithmic variables, so that, in the soft and collinear limit, equal areas correspond to equal emission probabilities.

We choose to draw Lund planes as depicted in Fig. 1. On the horizontal axis, we show the logarithm of the emission angle with respect to the jet axis, as measured in the rapidity-azimuth plane, normalised to the jet radius, i.e. $\theta = \frac{\Delta}{R_0}$. On the vertical axis, we show the emission's transverse momentum with respect to the jet axis, normalised to the hard scale

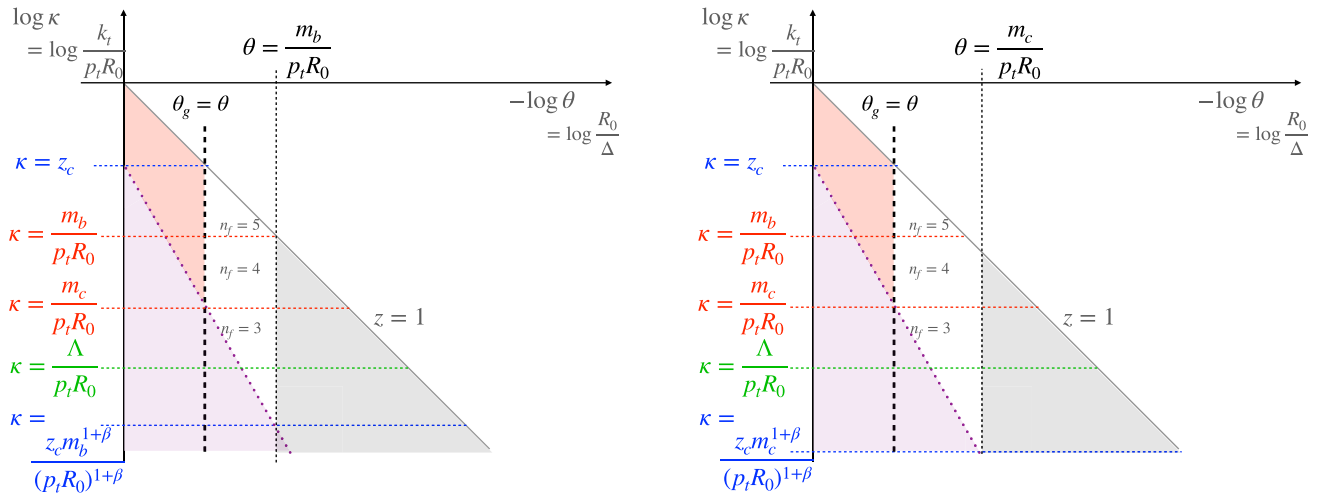

Fig. 1 Lund plane representation of the soft and quasi-collinear phase space for emissions off a b quark, on the left, and off a c quark, on the right. The dead-cone region is indicated in grey. Soft Drop is applied and the groomed-away region appears in light purple. The vertical dashed line in black indicates a measurement of the groomed radius θ_g and the corresponding area in pink is the vetoed region, which gives rise to the Sudakov form factor. The horizontal red lines indicate the boundaries between different flavour numbers for the running coupling, while the green one marks the boundary of the non-perturbative region

of the process, $\kappa = \frac{k_t}{p_t R_0}$. This way, lines of constant θ_g, in the soft and collinear limits, are represented by straight vertical lines, as shown in Fig. 1, in dashed black.

The Soft Drop condition $z = z_c\theta^\beta$ sets the dotted line in purple. In this study, we consider $\beta \geq 0$. The region of phase space below this line (shaded in purple) is groomed away.

Recently, two of us have achieved a generalisation of the Lund-plane resummation formalism that includes quark-mass effects [36]. The Lund plane on the left-hand panel of Fig. 1 is for a b-quark line, while the one on the right is for a c-quark. When considering heavy-quark Lund planes, two important differences with respect to the massless case arise. First, the presence of the vertical short-dashed line, in black, that represents the beginning of the dead cone at $\theta = \frac{m_b}{p_t R_0}$, i.e. $\Delta = \frac{m_b}{p_t}$, for the left-hand figure, and at $\theta = \frac{m_c}{p_t R_0}$, for the right-hand one. The shaded area in grey is the dead-cone region. Here, collinear emissions do not give rise to any logarithmic enhancement, leading to a suppression of QCD activity in this region. Second, the presence of heavy-quark thresholds, which are relevant when considering the running of the strong coupling. In Fig. 1, they are represented by horizontal dotted lines in red. The lines at $\kappa = \frac{m_{b,c}}{p_t R_0}$ correspond to $k_t = m_{b,c}$ and therefore mark the boundaries between two regions with different numbers of active flavours: $n_f = 5$ above and $n_f = 4$ below, and $n_f = 4$ above and $n_f = 3$ below, for $k_t = m_b$ and $k_t = m_c$, respectively. We also show (in green) the line $\kappa = \frac{\Lambda}{p_t R_0}$ that marks the $k_t = \Lambda \simeq 1$ GeV boundary between the perturbative and the non-perturbative regions. For the region below this line, we need a prescription to regulate the Landau pole of the strong coupling. Details are given in Appendix A. Finally, we show dotted lines in blue, at $\kappa = z_c$ and at $\kappa = z_c\left(\frac{m_{b,c}}{p_t R_0}\right)^{1+\beta}$, which originate from the Soft Drop condition and the dead-cone boundary.

When performing actual calculations, one needs to establish the hierarchy between these scales. Let us consider, as an example the case of b-jets, with relatively high-p_t. Because Soft Drop is usually employed with $z_c = 0.1$, we have that $z_c > \frac{m_b}{p_t R_0}$ if $p_t R_0 > 42$ GeV. Ideally, we would like to have $z_c\left(\frac{m_b}{p_t R_0}\right)^{1+\beta} > \frac{\Lambda}{p_t R_0}$. This results in an upper bound for the hard scale $(p_t R_0)^\beta < z_c \frac{m_b^{1+\beta}}{\Lambda}$. However, this inequality is hardly satisfied if we consider commonly used values of the Soft Drop parameters, e.g. $z_c = 0.1$ and $\beta = 0$. Thus, we do expect non-perturbative contributions to affect the dead-cone region and we will assess their size using Monte Carlo simulations. In Fig. 1, we show the Lund plane for this hierarchy of scales, on the left for b-jets and on the right for c-jets. A discussion of all the other possible hierarchies and regions can be found in Appendix B.

3.2 Recap of the massless calculation

Before computing the θ_g resummed distribution for a jet initiated by a massive quark, let us briefly recall the structure of the resummation in the case of massless partons. The all-order cumulative distribution can be written as the product of two contributions

$$\Sigma^{\mathrm{NLL}}(\theta_g^2) = S(\theta_g^2)\, e^{-R(\theta_g^2)}. \tag{4}$$

The function R is the Sudakov exponent, or radiator, which accounts for angular-ordered collinear emissions. It was first

computed in [37] and, for the case of a quark jet, reads

$$R(\theta_g^2) = \int_{\theta_g^2}^1 \frac{\mathrm{d}\theta^2}{\theta^2} \int_0^1 \mathrm{d}z\, P_{gq}(z) \frac{\alpha_{\mathrm{S}}^{\mathrm{CMW}}(k_t^2)}{2\pi} \Theta(z - z_c\theta^\beta), \tag{5}$$

where θ is the emission angle, normalised to the jet radius, $k_t = z\theta p_t R_0$ and the massless $q \to qg$ (timelike) splitting function is

$$P_{gq}(z) = C_{\mathrm{F}} \frac{1 + (1-z)^2}{z}. \tag{6}$$

In order to achieve NLL accuracy, the running coupling must be evaluated at two loops in the so-called CMW scheme [47]. Henceforth, we will also work in the small-z_c limit. Consequently, even in the case $\beta = 0$, we will ignore flavour-changing contributions [48,49]. The Sudakov form factor represents the no-emission probability and corresponds to the area shaded in pink on the Lund planes of Fig. 1.

This rather simple picture essentially arises from two facts. First, for the θ_g distribution, independent emissions, i.e. $\alpha_{\mathrm{S}}^n C_{\mathrm{F}}^n$ contributions, exponentiate with no corrections due to multiple emissions [37]. Second, all non-Abelian contributions are unresolved and therefore captured at NLL by the running coupling in the CMW scheme. However, as it was pointed out in [44], the above description is not complete at NLL and the resummed expression must be supplemented with a correction factor, S in Eq. (4). Indeed simple exponentiation is broken by single logarithmic contributions that arise from two competing mechanisms. Let us start with the correction to the independent emission contribution. The C/A algorithm, which is used as the first step of the Soft Drop procedure, can cluster two soft gluons together first, if they are the closest pair in angle, instead of clustering each of them with the hard quark. This happens when $\theta_{12} < \min(\theta_1, \theta_2)$. This effect introduces single-logarithmic corrections that are usually referred to as clustering logarithms [50–53]. On the other hand, the clustering condition also plays a role in the correlated emission contribution. The C/A algorithm can resolve a soft gluon splitting, giving rise to a non-global contribution [54,55] that spoils the CMW picture. This is also a single logarithmic contribution that can happen if the two soft gluons are not the closest pair in angle, i.e. $\theta_{12} > \min(\theta_1, \theta_2)$.

Both contributions start at $\mathcal{O}(\alpha_{\mathrm{S}}^2)$ and can be computed using the expression of the squared matrix element for the emission of two strongly-ordered soft gluons with momenta k_1, k_2 off a hard fermionic dipole with momenta p_a, p_b [56,57]

$$\begin{aligned} W &= 2C_{\mathrm{F}} w_{ab,1} \left[C_{\mathrm{A}} w_{a1,2} + C_{\mathrm{A}} w_{b1,2} + (2C_{\mathrm{F}} - C_{\mathrm{A}})\, w_{ab,2} \right] \\ &= 4C_{\mathrm{F}}^2\, w_{ab,1} w_{ab,2} + 2C_{\mathrm{F}} C_{\mathrm{A}}\, w_{ab,1} \left(w_{a1,2} + w_{b1,2} - w_{ab,2} \right), \end{aligned} \tag{7}$$

where the C_{F}^2 term describes the independent emission contribution, while the $C_{\mathrm{F}}C_{\mathrm{A}}$ one, the correlated one. We have introduced the eikonal factor

$$w_{ij,l} = \frac{p_i \cdot p_j}{p_i \cdot k_l\, p_j \cdot k_l}. \tag{8}$$

Following [44], we consider the case in which the softer gluon (k_2) is emitted at a large angle and passes the Soft Drop condition, while the harder gluon (k_1) is emitted at an angle smaller than θ_g and so it is not subject to the Soft Drop condition. If we work in the small-angle limit, considering real and virtual contributions, we find

$$\begin{aligned} S = 1 + \left(\frac{\alpha_{\mathrm{S}}}{\pi}\right)^2 \int_0^1 \frac{\mathrm{d}z_1}{z_1} \int_0^1 \frac{\mathrm{d}z_2}{z_2} \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \int_0^{\theta_g^2} \frac{\mathrm{d}\theta_1^2}{\theta_1^2} \int_{\theta_g^2}^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2} \\ \times \Theta\left(z_2 - z_c\theta_2^\beta\right) \Theta\left(z_1 - z_2\right) \Big[C_{\mathrm{F}}^2\, \Theta\left(\theta_1 - \theta_{12}\right) \\ - C_{\mathrm{F}} C_{\mathrm{A}} \frac{\theta_1\theta_2\, \cos\phi}{\theta_1^2 + \theta_2^2 - 2\theta_1\theta_2\cos\phi} \Theta\left(\theta_{12} - \theta_1\right) \Big] + \mathcal{O}(\alpha_{\mathrm{S}}^3), \end{aligned} \tag{9}$$

with $\theta_{12}^2 = \theta_1^2 + \theta_2^2 - 2\theta_1\theta_2\cos\phi$. The integration above can be simplified by noting that, in order to capture the highest power of the logarithm of θ_g, we can evaluate the momentum fraction integrals with lower limit $z_2 = z_c\theta_2^\beta \simeq z_c\theta_g^\beta$. This way, the momentum fraction integrals decouple from the angular ones and we obtain

$$\begin{aligned} S &= 1 + \frac{1}{2}\left(\frac{\alpha_{\mathrm{S}}}{\pi}\right)^2 \log^2\left(z_c\theta_g^\beta\right) \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \int_0^{\theta_g^2} \frac{\mathrm{d}\theta_1^2}{\theta_1^2} \int_{\theta_g^2}^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2} \\ &\quad \times \Bigg[C_{\mathrm{F}}^2\, \Theta\left(2\theta_1\cos\phi - \theta_2\right) - C_{\mathrm{F}} C_{\mathrm{A}} \frac{\theta_1\theta_2\, \cos\phi}{\theta_1^2 + \theta_2^2 - 2\theta_1\theta_2\cos\phi} \\ &\quad \times \Theta\left(\theta_2 - 2\theta_1\cos\phi\right) \Bigg] + \mathcal{O}\left(\alpha_{\mathrm{S}}^2 \log\theta_g\right) \\ &= 1 + \left(\frac{\alpha_{\mathrm{S}}}{\pi}\right)^2 \log^2\left(z_c\theta_g^\beta\right) \frac{\pi^2}{108} \left(C_{\mathrm{F}}^2 - 4C_{\mathrm{F}}C_{\mathrm{A}}\right). \end{aligned} \tag{10}$$

The all-order resummation of these contributions has been performed numerically, in the large-N_c limit [44]. In this study, we have decided to limit ourselves to studying the impact of these corrections. Thus, we approximate the clustering factor S by simply considering the exponentiation of the two-loop result:

$$S(\theta_g^2) = \exp\left[\frac{C_{\mathrm{F}}^2 - 4C_{\mathrm{F}}C_{\mathrm{A}}}{108} \left(\int_{z_c\theta_g^\beta}^1 \frac{\mathrm{d}z}{z}\, \alpha_{\mathrm{S}}(z^2 p_t^2 R_0^2) \right)^2 \right], \tag{11}$$

where the running of the strong coupling is taken at one loop.

3.3 The θ_g distribution for a heavy-flavour jet

We now perform the resummation of the θ_g distribution for jets initiated by a heavy quark, considering both the case of a b-quark and a c-quark. Both the Sudakov contribution R and the clustering correction function S in Eq. (4) need to be reconsidered. We start with the calculation of the resummed Sudakov exponent.

It was realised long ago that squared QCD matrix elements with massive partons factorise in the so-called quasi-collinear limit [58,59]. In this approximation, both the transverse momentum k_t of the emission with respect to the massive emitter, and the mass m are small compared to the hard scale but they are considered of the same order. In this limit, the squared invariant amplitude for one-gluon emission takes the form

$$|\mathscr{M}|^2 \simeq 8\pi\alpha_{\rm S}\frac{z(1-z)}{k_t^2+z^2m_i^2}P_{gi}\left(z,k_t^2\right)|\mathscr{M}_0|^2, \quad i=b,c, \tag{12}$$

where the massive splitting function for $i\to ig$ is

$$P_{gi}(z,k_t^2) = C_{\rm F}\left(\frac{1+(1-z)^2}{z}-\frac{2z(1-z)m_i^2}{k_t^2+z^2m_i^2}\right). \tag{13}$$

Following the same steps as in [36], we write

$$R_i(\theta_g^2,m_i^2,m_b^2,m_c^2) = \int_0^{p_t^2R_0^2}\frac{{\rm d}k_t^2}{k_t^2+z^2m_i^2}\int_0^1{\rm d}z\,P_{gi}(z,k_t^2) \times\frac{\alpha_{\rm S}^{\rm CMW}(k_t^2)}{2\pi}\Theta(z-z_c\theta^\beta)\Theta(\theta-\theta_g), \tag{14}$$

where $\theta=\frac{k_t}{zp_tR_0}$. Note that the mass dependence enters in two ways. First, for a jet initiated by a quark flavour $i=b,c$, the quasi-collinear phase-space and the massive splitting function depend on m_i. Second, regardless of the jet flavour, the running of the strong coupling may cross the b and the c thresholds, thus inducing a logarithmic dependence on the quark masses. In order to proceed further, it is convenient to change the integration variable in Eq. (14) from k_t^2 to θ^2:

$$R_i(\theta_g^2,\theta_i^2,\xi_b,\xi_c) = \int_{\theta_g^2}^1\frac{{\rm d}\theta^2}{\theta^2+\theta_i^2}\int_0^1{\rm d}z\,P_{gi}\left(z,k_t^2\right)\frac{\alpha_{\rm S}^{\rm CMW}(k_t^2)}{2\pi} \times\Theta(z-z_c\theta^\beta), \tag{15}$$

where now $k_t^2=z^2\theta^2p_t^2R_0^2$, and we have introduced the following dimensionless variables

$$\xi_i=\frac{m_i^2}{p_t^2R_0^2}, \quad \theta_i^2=\xi_i, \quad \text{with } i=b,c. \tag{16}$$

We stress again that the Sudakov exponent for either a high-p_t b-jet or c-jet depends on both ξ_b and ξ_c because the running coupling crosses both quark thresholds. However, it depends only on the dimensional ratio θ_i^2 of the corresponding flavour $i=b,c$.

The mass-dependent shift in the denominator of Eq. (15) acts as an effective lower bound of a logarithmic angular integration. This is the well-known dead-cone effect, [24,25] i.e. the fact that radiation off massive partons at angles below m/p_t is not logarithmically enhanced. Therefore, we can further simplify our expression by shifting the angular integration variable to $\bar\theta^2=\theta^2+\theta_i^2$. Neglecting power corrections in the mass, we have

$$R_i(\theta_g^2,\theta_i^2,\xi_b,\xi_c) = \int_{\vartheta_{g,i}^2}^1\frac{{\rm d}\bar\theta^2}{\bar\theta^2}\int_0^1{\rm d}z\,P_{gi}\left(z,z^2(\bar\theta^2-\theta_i^2)p_t^2R_0^2\right) \times\frac{\alpha_{\rm S}^{\rm CMW}(z^2(\bar\theta^2-\theta_i^2)p_t^2R_0^2)}{2\pi}\Theta(z-z_c(\bar\theta^2-\theta_i^2)^{\frac{\beta}{2}}), \tag{17}$$

with $\vartheta_{g,i}^2=\theta_g^2+\theta_i^2$. Let us first note that

$$P_{gi}\left(z,z^2(\bar\theta^2-\theta_i^2)p_t^2R_0^2\right) = C_{\rm F}\left(\frac{1+(1-z)^2}{z}-\frac{2\theta_i^2(1-z)}{z\bar\theta^2}\right) \equiv\mathscr{P}_{gi}(z,\bar\theta^2). \tag{18}$$

Furthermore, as discussed in detail in [36], the mass-dependent shift in the argument of the running coupling contributes at most to NNLL corrections and, therefore, it can be dropped at the accuracy we are working at. A similar argument applies to the shift in the Soft Drop condition. This can be easily checked at fixed coupling. For instance, in the soft and collinear limits, we have

$$\begin{aligned}R_i^{\rm (f.c.)} &= \frac{\alpha_{\rm S}C_{\rm F}}{\pi}\int_{\vartheta_{g,i}^2}^1\frac{{\rm d}\bar\theta^2}{\bar\theta^2}\int_0^1\frac{{\rm d}z}{z}\Theta(z-z_c(\bar\theta^2-\theta_i^2)^{\frac{\beta}{2}})\\ &= \frac{\alpha_{\rm S}C_{\rm F}}{\pi}\left\{\log\vartheta_{g,i}^2\left(\log z_c+\frac{\beta}{4}\log\vartheta_{g,i}^2\right)\right.\\ &\quad\left.-\frac{\beta}{2}\left[{\rm Li}_2\left(-\frac{1}{\theta_i^2}\right)-{\rm Li}_2\left(\frac{\theta_i^2}{\vartheta_{g,i}^2}\right)+\frac{1}{2}\log^2\theta_i^2+\frac{\pi^2}{6}\right]\right\}\\ &= \frac{\alpha_{\rm S}C_{\rm F}}{\pi}\int_{\vartheta_{g,i}^2}^1\frac{{\rm d}\bar\theta^2}{\bar\theta^2}\int_0^1\frac{{\rm d}z}{z}\Theta(z-z_c\bar\theta^\beta)+{\rm NNLL},\end{aligned} \tag{19}$$

where in the last step we have dropped constant terms in the $\theta_i\to0$ or $\theta_g\to0$ limits, which are NNLL. Thus, the massive Sudakov exponent at NLL becomes

$$R_i(\theta_g^2,\theta_i^2,\xi_b,\xi_c) = \int_{\vartheta_{g,i}^2}^1\frac{{\rm d}\bar\theta^2}{\bar\theta^2}\int_{z_c\bar\theta^\beta}^1{\rm d}z\,\mathscr{P}_{gi}(z,\bar\theta^2) \times\frac{\alpha_{\rm S}^{\rm CMW}(z^2\bar\theta^2p_t^2R_0^2)}{2\pi}. \tag{20}$$

Remarkably, with the sole exception of the mass-dependent term in the splitting function, this result, when seen as a function of $\vartheta^2_{g,i}$, has the same form of the massless case, Eq. (5). Consequently, we can represent the massive Sudakov form factor in Fig. 1 using the same Lund planes as in the massless case, provided that we interpret vertical lines as lines of constant $\vartheta^2_{g,i}$. This way, the dead-cone line acts as a phase-space boundary because $\vartheta^2_{g,i} \to \theta_i^2$ as $\theta_g \to 0$. The integrations in Eq. (20) are all straightforward but the presence of the Soft Drop condition, of the dead cone, and of mass thresholds force us to consider many cases. Details are given in the Appendices.

Next, we have to consider the clustering correction factor S. We adopt the same, approximate, strategy of the massless case. Namely, we consider the running-coupling exponentiation of the two-loop result. The square amplitude for the emission of two soft gluons off a massive quark–antiquark dipole was computed in [60,61]. For our purposes, we are actually interested in a less general result, namely the case of the emission of two soft gluons that are strongly ordered in energies, e.g. $z_2 \ll z_1$, off a single dipole. In this limit, the massive square matrix element takes the same form as Eq. (7) but with the massless eikonal factor $w_{ij,k}$ substituted by the massive one:

$$w^m_{ij,l} = \frac{p_i \cdot p_j}{p_i \cdot k_l \, p_j \cdot k_l} - \frac{m_i^2}{2p_i \cdot k_l} - \frac{m_j^2}{2p_j \cdot k_l}. \tag{21}$$

We choose p_a, see Eq. (7), to be the momentum of the heavy quark and we work in the quasi-collinear limit with respect to its momentum direction. Therefore, we can set $p_b^2 = 0$. We then integrate the C_F^2 and $C_\mathrm{F}C_\mathrm{A}$ contributions with the same constraints as in Eq. (9):

$$\begin{aligned} S_i = 1 + \left(\frac{\alpha_\mathrm{S}}{\pi}\right)^2 \int_0^1 \frac{\mathrm{d}z_1}{z_1} \int_0^1 \frac{\mathrm{d}z_2}{z_2} \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \int_0^{\theta_g^2} \frac{\mathrm{d}\theta_1^2}{\bar\theta_1^2} \int_{\theta_g^2}^1 \frac{\mathrm{d}\theta_2^2}{\bar\theta_2^2} \\ \times \Theta\left(z_2 - z_c\theta_2^\beta\right) \Theta\left(z_1 - z_2\right) \Bigg[C_\mathrm{F}^2 \, \Theta\left(\theta_1 - \theta_{12}\right) \\ \times \left(1 - \frac{\theta_i^2}{\bar\theta_1^2} - \frac{\theta_i^2}{\bar\theta_2^2} + \frac{\theta_i^4}{\bar\theta_1^2\bar\theta_2^2}\right) - C_\mathrm{F}C_\mathrm{A} \frac{\theta_1\theta_2 \, \cos\phi + \theta_i^2}{\theta_1^2 + \theta_2^2 - 2\theta_1\theta_2\cos\phi} \\ \times \left(1 - \frac{\theta_i^2}{\bar\theta_1^2}\right) \Theta\left(\theta_{12} - \theta_1\right) \Bigg] + \mathscr{O}(\alpha_\mathrm{S}^3), \quad i = b, c, \end{aligned} \tag{22}$$

where, as before, $\bar\theta_1^2 = \theta_1^2 + \theta_i^2$ and $\bar\theta_2^2 = \theta_2^2 + \theta_i^2$. As in the massless case, in order to obtain the leading contribution we can substitute $\theta_2 \simeq \theta_g$ in the Soft Drop condition. Therefore, the momentum fraction integrals decouple from the angular ones and we obtain

$$S_i = 1 + \left(\frac{\alpha_\mathrm{S}}{\pi}\right)^2 \log^2 z_c\theta_g^\beta \Big[C_\mathrm{F}^2 \mathscr{F}_1(\theta_g^2, \theta_i^2) + C_\mathrm{F}C_\mathrm{A} \mathscr{F}_2(\theta_g^2, \theta_i^2) \Big]. \tag{23}$$

The Abelian contribution reads

$$\begin{aligned} \mathscr{F}_1(\theta_g^2, \theta_i^2) = \int_0^{\frac{\pi}{3}} \frac{\mathrm{d}\phi}{2\pi} \int_1^{4\cos^2\phi} \frac{x \, \mathrm{d}x}{(x+\kappa)^2} \Bigg[\log\left(g(\phi)\right) \\ - \frac{\kappa}{1+\kappa}\left(g(\phi) - 1\right) \Bigg], \end{aligned} \tag{24}$$

with

$$g(\phi) = \frac{4(1+\kappa)\cos^2\phi}{x + 4\kappa\cos^2\phi}, \quad \kappa = \frac{\theta_i^2}{\theta_g^2}. \tag{25}$$

For the non-Abelian case, instead, we find

$$\begin{aligned} \mathscr{F}_2(\theta_g^2, \theta_i^2) = -\int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \int_0^1 \frac{\mathrm{d}y}{y(1+y^2\kappa)} \int_0^y \mathrm{d}t \frac{2t}{t^2+y^2\kappa} \\ \times \frac{t\cos\phi + y^2\kappa}{t^2 - 2t\cos\phi + 1} \left[1 - \frac{\kappa y^2}{t^2 + \kappa y^2}\right] \Theta(1 - 2t\cos\phi). \end{aligned} \tag{26}$$

We evaluate $\mathscr{F}_1$ and $\mathscr{F}_2$ numerically. However, in the limit $\theta_g^2 \gg \theta_i^2$ we are able to perform the integrals analytically and we obtain

$$\begin{aligned} \lim_{\theta_i^2/\theta_g^2 \to 0} \mathscr{F}_1(\theta_g^2, \theta_i^2) &= \frac{\pi^2}{108}, \\ \lim_{\theta_i^2/\theta_g^2 \to 0} \mathscr{F}_2(\theta_g^2, \theta_i^2) &= -\frac{\pi^2}{27}, \end{aligned} \tag{27}$$

so that the massless result Eq. (10) is recovered. Remarkably, in the opposite limit, we find

$$\lim_{\theta_g^2/\theta_i^2 \to 0} \mathscr{F}_1(\theta_g^2, \theta_i^2) = \lim_{\theta_g^2/\theta_i^2 \to 0} \mathscr{F}_2(\theta_g^2, \theta_i^2) = 0, \tag{28}$$

so that these contributions disappear asymptotically. We note that this result is not related to the clustering conditions but rather to the requirement that one of the two gluons (k_1) be within the groomed jet radius, i.e. $0 < \theta_1^2 < \theta_g^2$. Changing the integration variable to $\bar\theta_1^2$, we have $\theta_i^2 < \bar\theta_1^2 < \theta_g^2 + \theta_i^2$. Thus, we have to evaluate the integral of a (regular or integrable) function over a domain that has a vanishing measure as $\theta_g^2 \to 0$.

As for the massless case, we do not perform the full NLL resummation of these effects, but we limit ourselves to exponentiate the $\mathscr{O}(\alpha_\mathrm{S}^2)$ with running coupling corrections, as done in Eq. (11):

$$\begin{aligned} S_i(\theta_g^2, \theta_i^2, \xi_b, \xi_c) = \exp\Bigg[\frac{C_\mathrm{F}^2 \mathscr{F}_1(\theta_g^2, \theta_i^2) + C_\mathrm{F}C_\mathrm{A} \mathscr{F}_2(\theta_g^2, \theta_i^2)}{\pi^2} \\ \times \left(\int_{z_c\theta_g^\beta}^1 \frac{\mathrm{d}z}{z} \, \alpha_\mathrm{S}(z^2 p_t^2 R_0^2) \right)^2 \Bigg]. \end{aligned} \tag{29}$$

In summary, the resummed θ_g distribution for a heavy-flavour jet reads

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i}{\mathrm{d}\theta_g} = \frac{\mathrm{d}}{\mathrm{d}\theta_g}\left[S_i(\theta_g^2, \theta_i^2, \xi_b, \xi_c)\, e^{-R_i(\theta_g^2,\theta_i^2,\xi_b,\xi_c)}\right], \tag{30}$$

for $i = b, c$. The Sudakov exponent R_i is given in Eq. (20) and the clustering contribution S_i in Eq. (29). Note that our NLL distribution is no longer normalised, as a consequence of the fact that, in the massive case, the resummed cumulative behaves as a constant for $\theta_g \to 0$, while it vanishes in the massless case, see Eq. (4). Indeed, while in the massless case, Soft Drop with $\beta \geq 0$ behaves perturbatively as a groomer, i.e. within resummed perturbation theory, it always returns a jet with $\theta_g > 0$, the quark mass provides an effective cutoff so that, there is a non-vanishing probability, given by

$$S_i(0, \theta_i^2, \xi_b, \xi_c)\, e^{-R_i(0,\theta_i^2,\xi_b,\xi_c)} \sim e^{-\frac{\alpha_S C_F}{\pi\beta}\left(\log^2(z_c\theta_i^\beta)-\log^2(z_c)\right)}, \tag{31}$$

to find Soft Drop jets with $\theta_g = 0$, even if we expect fixed-order corrections, as well as non-perturbative effects, to smear this effect out. Thus, we could either use Soft Drop in grooming mode and supplement Eq. (30) with an endpoint, $\delta(\theta_g)$, contribution that ensures normalisation or consider the algorithm in tagging mode and discard jets that return $\theta_g = 0$. Henceforth, we choose this second option.

Finally, we note that, in the calculations of the resummed exponent presented so far, we have always included virtual corrections, as well as the contributions from jets that fail Soft Drop. In tagging mode, one should discard those and, in principle, repeat the calculations. However, we observe that in resummed perturbation theory

$$\Sigma(\theta_g)\Big|_{\text{tagging}} = \Sigma(\theta_g)\Big|_{\text{grooming}} - \Sigma(0)\Big|_{\text{grooming}}. \tag{32}$$

Thus, in practice, we can just use the results obtained so far, provided that we remove the $\theta_g = 0$ contribution. Numerical results and their comparison to Monte Carlo simulations will be presented in the next section.

3.4 Numerical results and comparison to Monte Carlo simulations

We now provide numerical results for the θ_g distribution according to Eq. (30), and compare them to Monte Carlo (MC) simulations. To this purpose, we generate events using HERWIG, version 7.2.2 [62]. To test the all-order behaviour of the observables we are interested in, we consider LO matrix elements for the hard process, dressed with the parton shower.[1] Hadronisation effects are included using the HERWIG cluster hadronisation model, when explicitly declared, and the CT14 [66] set of PDFs has been used throughout the paper.

We simulate the inclusive production of a pair of oppositely charged muons in association with a jet, in proton–proton collisions at 13 TeV centre-of-mass energy. The muon pair is required to have an invariant mass between 70 and 110 GeV. Jets are clustered using the anti-k_t algorithm [41] with $R_0 = 0.4$ and then ordered in transverse momentum. The hardest jet containing a b (c) quark, or a B (D) hadron, is considered. To match the heavy quark/hadron with a jet we look at the closest (with respect to the jet axis) flavoured particle, starting from the highest in transverse momentum, with $p_t > 5$ GeV. Properties of the flavoured particle and the jet are extracted and analysed using RIVET [67] and FASTJET [68]. The fiducial phase-space for the muons is defined by the following cuts: $p_{t,\mu} > 26$ GeV, $|\eta_\mu| < 2.4$, while the jets are selected in the region defined by $|\eta_J| < 2.4$. We consider three different transverse-momentum regions, i.e. $p_t \geq 50$, 150 and 300 GeV. However, only $p_t \geq 150$ GeV is shown in the main text, and the other cases can be found in Appendix C.

The θ_g distribution is shown in Fig. 2. Each plot includes parton level, i.e. with parton shower effect only, and hadron level, i.e. with hadronisation and the Underlying Event (UE) included. Together with the MC prediction, we show the corresponding NLL result of Eq. (30). The latter also exhibits a theoretical uncertainty given by the variation of the renormalisation scale, for which the central value is set at the hard scale $p_t R_0$, by a factor of two, as customary. Plots in different rows correspond to results for the identified leading jet in $Z+b$ (B), $Z+c$ (D) and $Z+$ light quark/hadron production, respectively. We always consider Soft Drop with $z_c = 0.1$, and $\beta = 0$ and 1, for left and right plots, respectively. For b and c jets, we also indicate, in green, the expected dead-cone region $\theta_g < \theta_i = \frac{m_i}{p_t R_0}$. All curves are normalised to have unit area. For $\beta = 0$, i.e. when Soft Drop is more aggressive, we find good agreement between the NLL prediction and the MC. They are fairly different, instead, for $\beta = 1$ at high θ_g. However, in this region, we expect fixed-order corrections, not considered here, to be important.

In order to highlight possible dead-cone effects, it is useful to consider the ratio between the heavy-flavour jet θ_g distribution and the corresponding one for light-quark jets. We do this in Fig. 3. We concentrate on b-jets, but we show the results for three different transverse momentum cuts, namely 50, 150 and 300 GeV (from top to bottom). We show results for both $\beta = 0$ (left) and $\beta = 1$ (right). In every plot, we

[1] In this paper we concentrate on an angular-ordered parton shower, see [63] for detailed comparisons with analytic resummation. We have also tested our analytic predictions against the HERWIG dipole shower [64] and the PYTHIA8.3 [65] one and we did not find significant differences. It would be interesting, in the future, to also compare to state-of-the-art simulations that also include matching to fixed order.

show the ratio obtained with our NLL resummed prediction and with the simulation performed with HERWIG, both at parton-level and hadron-level.

As expected, because $\theta_i = \frac{m_i}{p_t R_0}$, the dead cone is more visible at lower values of p_t. Interestingly, deviations between the b and the light quark distribution start at angular scales bigger than θ_i. However, as p_t is increased, this transition is pushed to a region that is likely beyond experimental resolution. We also note that mass effects are more pronounced for the $\beta = 0$ case than $\beta = 1$. Considering that we have already noted that the former is under better theoretical control than the latter, the Soft Drop jets with $\beta = 0$ appear to be an interesting choice to study the dead cone.

4 The z_g distribution

We now discuss the second variable that characterises Soft Drop jets, namely the momentum fraction z_g. In this discussion, we want to provide a simple theoretical prediction for z_g for jets initiated by massive quarks. We also compare, both analytically and in Monte Carlo simulation, the variable z_g to a widely-used fragmentation variable that measures the transverse momentum fraction of the $B(D)$ hadrons (or $b(c)$ quark at parton level) with respect to the jet p_t.

4.1 Recap of the massless calculation

We start by briefly reviewing the calculation of the z_g distribution for light jets. The value of z_g is fixed by the first de-clustering of the jet that passes the Soft Drop condition. Because we are completely inclusive over the splitting angle, we must integrate over all possible values of θ_g, including configurations where the two emissions become collinear, for which the integral diverges. If $\beta \geq 0$, collinear splittings always pass the Soft Drop condition and these divergent configurations are not cancelled by the corresponding virtual corrections, for which z_g is undefined, and heralds the fact that the observable is not IRC safe.

However, z_g belongs to a wider class, i.e. Sudakov safe observables [37,38,69,70], that despite being IRC unsafe, can be computed in perturbation theory, provided that we use resummation. For this purpose, we need to introduce a safe companion observable. The Soft Drop procedure itself suggests using the groomed angle θ_g, which we have discussed in the previous section. Therefore, we imagine to measure a value of z_g, given a finite angular separation θ_g. Using the language of conditional probabilities, we have [38]:

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma}{\mathrm{d}z_g} = \int_0^1 d\theta_g \; p(\theta_g)\, p(z_g|\theta_g), \tag{33}$$

where $p(\theta_g) = \frac{1}{\sigma_0}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}$ and $p(z_g|\theta_g) = \frac{p(z_g,\theta_g)}{p(\theta_g)}$ is the conditional probability for measuring z_g given a value of θ_g. If $\beta < 0$, z_g is IRC safe and the integral in Eq. (33) can be computed order by order in α_S. This is no longer true when $\beta \geq 0$, which is the standard configuration in which the Soft Drop algorithm is used and, therefore, our case of interest. In this situation, the integral (33) diverges order by order in the strong coupling because of the $1/\theta_g$ behaviour of the integrand. However, if we take $p(\theta_g)$ to be the resummed distribution, i.e. the derivative of Eq. (4), then the Sudakov form factor regulates the $\theta_g = 0$ singularity, providing us with a finite answer for Eq. (33).

Because in this work we are not concerned with logarithms of z_g and z_c, we evaluate the conditional probability at fixed-order, focusing on the case of a quark-initiated jet[2]:

$$p(z_g|\theta_g) = \frac{P_{gq}(z_g)\alpha_S(z_g^2\theta_g^2 p_t^2 R_0^2)}{\int_{z_c\theta_g^\beta}^1 \mathrm{d}z\, P_{gq}(z)\alpha_S(z^2\,\theta_g^2 p_t^2 R_0^2)}\Theta(z_g - z_c\theta_g^\beta). \tag{34}$$

The Soft Drop condition (1) dictates $z_g < \frac{1}{2}$ and, if $\beta = 0$, $z_g > z_c$.

The integral in Eq. (33) with running coupling must be performed numerically. However, it is interesting to study its fixed-coupling and lowest-order approximation, see Eq. (19). Thus, we consider

$$\begin{aligned} R^{(\mathrm{f.c.})}(\theta_g^2) &= \frac{\alpha_S C_F}{\pi\beta}\Big(\log^2(z_c\theta_g^\beta) - \log^2(z_c)\Big) \\ &= \frac{\alpha_S C_F}{\pi}\Big(\frac{\beta}{4}\log^2\theta_g^2 + \log\theta_g^2\,\log z_c\Big), \\ S^{(\mathrm{f.c.})}(\theta_g^2) &= 1. \end{aligned} \tag{35}$$

For $\beta \geq 0$, we find [38]

$$\begin{aligned} \frac{1}{\sigma_0}\frac{\mathrm{d}\sigma^{(\mathrm{f.c.})}}{\mathrm{d}z_g} &= P_{gq}(z_g)\frac{\alpha_S}{2\pi}\int_0^1 \frac{\mathrm{d}\theta_g^2}{\theta_g^2} e^{-R(\theta_g^2)}\Theta(z_g - z_c\theta_g^\beta) \\ &= \sqrt{\frac{\alpha_S}{4\beta C_F}}P_{gq}(z_g)e^{\frac{\alpha_S C_F}{\pi\beta}\log^2 z_c}\left[\mathrm{erf}\left(\sqrt{\frac{\alpha_S C_F}{\pi\beta}}\log a\right) + 1\right], \end{aligned} \tag{36}$$

[2] We note that the formalism we have just presented does not describe the situation in which we have a $q \to qg$ splitting, with the gluon harder than the quark, i.e. when the minimum function of the Soft Drop condition (1) returns $1-z$. This case can be accounted for by introducing a symmetrised version of the splitting function [38,71,72], which is integrated over $0 < z < 1/2$. For consistency, one should then make this replacement also in the calculation of the resummed θ_g distribution. In what follows, in order to streamline our discussion, we keep the standard version of the splitting function.

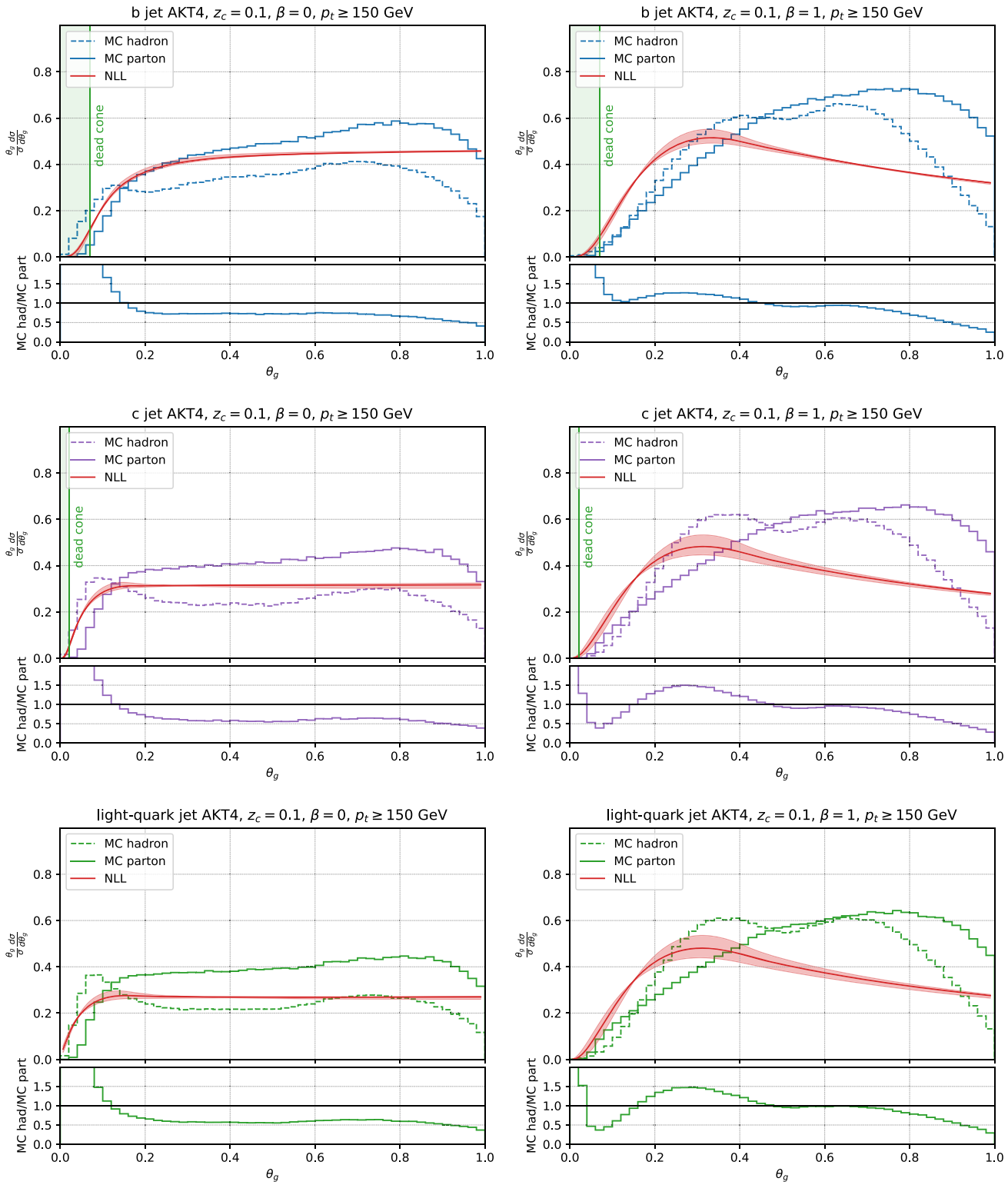

Fig. 2 The groomed jet radius (θ_g) distribution for b-jets (top), c-jets (middle) and light-quark jets (bottom). In all plots, we show our NLL resummation, as well as the results obtained with the Monte Carlo event generator HERWIG, both at parton and hadron level. Jets are selected with the anti-k_t algorithm with $R_0 = 0.4$, with $p_t \geq 150$ GeV, and groomed with Soft Drop with $z_c = 0.1$, and two values of the angular exponent: $\beta = 0$ (on the left) and $\beta = 1$ (on the right). The uncertainty bands for the analytic predictions are obtained by varying the resummation scale by a factor of two above and below the hard scale $p_t R_0$, i.e. $\mu_R \in \left[\frac{p_t R_0}{2}, 2 p_t R_0\right]$

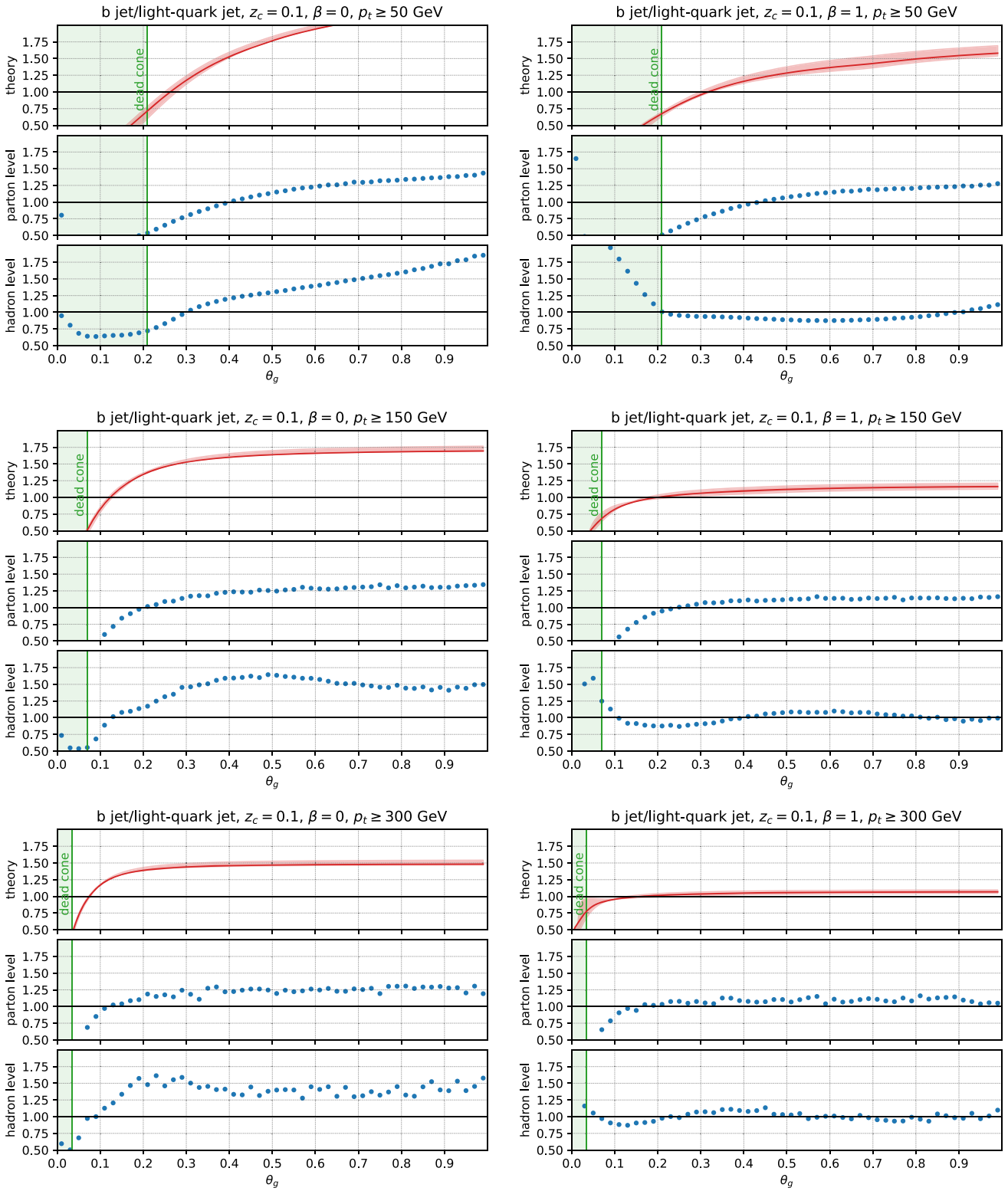

Fig. 3 Ratios of groomed jet radius (θ_g) distribution of b jets to light-quark jets for three values of the jet transverse momentum: 50 GeV (top), 150 GeV (middle) and 300 GeV (bottom). The plots on the left are for $\beta = 0$, and the ones on the right are for $\beta = 1$. In each plot, we show our NLL resummation and the results obtained with HERWIG, both a parton-level and hadron-level

where $\mathrm{erf}(x)$ is the error function and $a = \min[z_g, z_c]$. If $\beta > 0$, the above result is non-analytic in α_S:

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma^{(\mathrm{f.c.})}}{\mathrm{d}z_g} = \frac{1}{2}\sqrt{\frac{\alpha_S}{\beta C_F}}\,P_{gq}(z_g) + \mathscr{O}(\alpha_s)\,. \tag{37}$$

Even more interesting is the $\beta = 0$ case. At fixed coupling, the conditional probability, Eq. (34), becomes independent of θ_g and factors out of the integration to give

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma^{(\mathrm{f.c.})}}{\mathrm{d}z_g} = \frac{P_{gq}(z_g)}{\int_{z_c}^1 \mathrm{d}z\, P_{gq}(z)}\Theta(z_g > z_c) + \mathscr{O}(\alpha_s)\,. \tag{38}$$

It can be shown that the $\beta = 0$ case does have a valid perturbative expansion in α_S, despite being α_S-independent at lowest order. Eq. (38) shows that the $\beta = 0$ z_g-distribution is essentially driven by the QCD splitting function [38]. This observation has initiated numerous theoretical [73–76] and experimental [77–83] studies (see also [71,72]) that aim to use z_g as a probe of QCD dynamics, both in pp and heavy-ion collisions.

4.2 The z_g distribution for a heavy-flavour jet

We now move to the case of b and c jets and we start by considering the simple $\mathscr{O}(\alpha_S)$ calculation of the z_g distribution, in the quasi-collinear limit:

$$\begin{aligned}\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.o.})}}{\mathrm{d}z_g} &= \frac{\alpha_S}{2\pi}\int_0^1 \frac{\mathrm{d}\theta^2}{\theta^2+\theta_i^2}\int_{z_c\theta^\beta}^1 \mathrm{d}z\, P_{gi}\left(z, (z\theta p_t R_0)^2\right)\\ &\quad\times\delta(z - z_g)\\ &= \frac{\alpha_S}{2\pi}\int_{\theta_i^2}^1 \frac{\mathrm{d}\bar\theta^2}{\bar\theta^2}\,\mathscr{P}_{gi}\left(z_g, \bar\theta^2\right)\Theta(z_g - z_c\bar\theta^\beta)\\ &\quad+\mathrm{NNLL},\end{aligned} \tag{39}$$

where, as before, we have dropped the mass-dependent shift in the Soft Drop condition. Even before performing the integral, it is clear that the mass of the heavy quark, as one might have expected, regulates the collinear singularity and so the z_g distribution is IRC safe, for every value of β. The computation of the integral is straightforward but the presence of the $\bar\theta^2$ dependent contribution in the splitting function complicates the result. However, to illustrate our point, it is sufficient to work at LL. Therefore, we approximate $\mathscr{P}_{gi} = 2C_F/z_g$, and we find

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.o.})}}{\mathrm{d}z_g} = -\frac{\alpha_S C_F}{\pi}\frac{1}{z_g}\begin{cases}\log\theta_i^2 & z_g > z_c,\\ \frac{2}{\beta}\log\frac{z_c\theta_i^\beta}{z_g} & z_c\theta_i^\beta < z_g < z_c.\end{cases} \tag{40}$$

Note that in the case $\beta = 0$, we have $z_g > z_c$ and only the first term survives. Albeit finite, this expression contains logarithms of θ_i^2, which become large in the boosted regime. The all-order resummation of logarithms of z_g partially addresses this problem. Indeed, keeping our focus on the LL fixed-coupling approximation, we find the following result for the normalised cumulative distribution:

$$\begin{aligned}\log\Sigma_i^{(\mathrm{f.c.})}(z_g) = \frac{\alpha_S C_F}{\pi}\Bigg[&-\log\theta_i^2\,\log z_g\,\Theta(z_g - z_c)\\ &+\left(\frac{1}{\beta}\log^2\frac{z_g}{z_c} - \log\theta_i^2\,\log z_g\right)\Theta(z_c - z_g)\Theta(z_g - z_c\theta_i^\beta)\\ &+\left(\frac{1}{\beta}\log^2\theta_i^\beta - \log\theta_i^2\,\log z_c\theta_i^\beta\right)\Theta(z_c\theta_i^\beta - z_g)\Bigg].\end{aligned} \tag{41}$$

The $\beta = 0$ case is rather simple

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.c.})}}{\mathrm{d}z_g} = -\frac{\alpha_S C_F}{\pi}\frac{\log\theta_i^2}{z_g}e^{-\frac{\alpha_S C_F}{\pi}\log\theta_i^2\,\log z_g}, \quad z_g > z_c. \tag{42}$$

We note that these expressions indeed resum those logarithms of θ_i that are associated with logarithms of z_g. However, the $\theta_i \to 0$ limit is not smooth and we do not recover the massless result of Eq. (36). This is related to the non-commutativity of the soft and massless limits, discussed at length in [35,36].

Another way of resumming logarithms θ_i^2 is to resort to the conditional probability procedure described above for the massless case. To illustrate the procedure, we repeat, for the massive case, the calculation that led to the fixed-coupling result in Eq. (36). At LL, the splitting function can be approximated by its soft contribution. Moreover, at this accuracy $R_i(\theta_g^2, \theta_i^2, \xi_b, \xi_c) = R(\vartheta_{g,i}^2)$, with R given by Eq. (35) and $\vartheta_{g,i}^2 \in [\theta_i^2, 1]$. Therefore, for $\beta \geq 0$, we find

$$\begin{aligned}\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.c.})}}{\mathrm{d}z_g} &= \frac{\alpha_S C_F}{\pi}\frac{1}{z_g}\int_{\theta_i^2}^1 \frac{\mathrm{d}\bar\theta^2}{\bar\theta^2}e^{-R(\bar\theta^2)}\Theta\left(z_g - z_c\bar\theta^\beta\right)\\ &= \sqrt{\frac{\alpha_S C_F}{\beta}}\frac{1}{z_g}e^{\frac{\alpha_S C_F}{\pi\beta}\log^2 z_c}\\ &\quad\times\left[\mathrm{erf}\left(\sqrt{\frac{\alpha_S C_F}{\pi\beta}}\log a\right) - \mathrm{erf}\left(\sqrt{\frac{\alpha_S C_F}{\pi\beta}}\log\left(z_c\theta_i^\beta\right)\right)\right],\end{aligned} \tag{43}$$

with $a = \min[z_g, z_c]$. In particular, for $\beta = 0$, we find

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.c.})}}{\mathrm{d}z_g} = \frac{1}{z_g\log z_c}\left(e^{-\frac{\alpha_S C_F}{\pi}\log\theta_i^2\log z_c} - 1\right), \quad z_g > z_c. \tag{44}$$

The first-order expansions of the above expressions agree with Eq. (40) and large logarithmic corrections in θ_i^2 are resummed. Furthermore, if we take $\theta_i \to 0$, we recover the massless distribution of Eq. (36). We further note that

Eq. (44) is the same as its massless counterpart, but for the normalisation factor in brackets. This makes sense because in the LL fixed-coupling approximation, both distributions are driven by the most singular part of the splitting function, which is the same for P_{gq}, P_{gb}, and P_{gc}.

Because it has the correct $\theta_i \to 0$ limit, we decide to use the conditional-probability approach to compute the z_g distribution of a heavy-flavour jet:

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i}{\mathrm{d}z_g} = \frac{1}{\sigma_0}\int_{\theta_i^2}^1 \mathrm{d}\bar{\theta}^2 \frac{\mathscr{P}_{gi}(z_g,\bar{\theta}^2)\alpha_{\mathrm{S}}(z_g^2\bar{\theta}^2 p_t^2 R_0^2)}{\int_{z_c\bar{\theta}^\beta}^1 dz\, \mathscr{P}_{gi}(z,\bar{\theta}^2)\alpha_{\mathrm{S}}(z^2\bar{\theta}^2 p_t^2 R_0^2)} \times \Theta(z_g - z_c\bar{\theta}^\beta)\left.\frac{\mathrm{d}\sigma_i}{\mathrm{d}\theta_g^2}\right|_{\theta_g^2=\bar{\theta}^2-\theta_i^2}, \quad i=b,c, \tag{45}$$

where the resummed θ_g distribution is given in Eq. (45). We can make some further simplifications. The derivative in Eq. (45) gives a factor that simplifies, within our accuracy, the denominator of the conditional probability. Thus, we obtain

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i}{\mathrm{d}z_g} = \frac{1}{\sigma_0}\int_{\theta_i^2}^1 \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}\mathscr{P}_{gi}(z_g,\bar{\theta}^2)\alpha_{\mathrm{S}}(z_g^2\bar{\theta}^2 p_t^2 R_0^2) \times S_i(\bar{\theta}^2,\theta_i^2,\xi_b,\xi_c)\, e^{-R_i(\bar{\theta}^2,\theta_i^2,\xi_b,\xi_c)}\Theta(z_g - z_c\bar{\theta}^\beta). \tag{46}$$

Numerical results and their comparison to Monte Carlo simulations will be presented in the next section.

We close this discussion by noting that Eq. (45) does not systematically resum logarithms of z_g and z_c. This is acceptable for our purposes because we are mostly interested in the $\beta = 0$ case, for which $z_g > z_c = 0.1$. However, it would be interesting to extend the full NLL resummation for z_g, performed for light jets in [45] to the massive case. Because the calculation in [45] is based on the resummation of the double differential (θ_g, z_g), it may overcome the difficulties about the $\theta_i \to 0$ limit of Eqs. (41) and (42), previously discussed.

4.3 Numerical results and comparison to Monte Carlo simulations

In this section, we compare our resummed results for the z_g distribution to the ones obtained with Monte Carlo event generators, both at parton- and hadron-level. We limit ourselves to the case $\beta = 0$ and we simulate events using Herwig, with the same settings as the ones described in Sect. 3.4.

In Fig. 4, we compare resummed results for two different values of the jet transverse momentum cut, namely $p_t \geq 50$ GeV, on the left, and $p_t \geq 150$ GeV, on the right. The plots at the top are for b-jets, the ones in the middle for c-jets, and the ones at the bottom for light-quark jets. All curves are normalised to have unit area. As already pointed out, the results all look very similar, because the shape of the distribution is mostly driven by the singular part of the splitting function for the emission of a gluon off a (massive) quark. However, we also note that differences between b,c, and light flavours are larger in the Monte Carlo results than in the analytic ones. This could be due to quark masses influencing the kinematics, thus causing mass-dependent recoil effects. These are, to a certain extent, accounted for in the parton shower, but neglected in the analytic resummation.

Overall, we find good agreement between our calculation and the Monte Carlo results, although our results are strangely closer to the full simulation than to the parton-level one. We note that our predictions undershoot the Monte Carlo at large z_g. We have traced this back to the fact that our calculation does not take into account symmetrised splitting functions, as discussed in the footnote of Sect. 4.1. Indeed by rescaling the z_g distribution by the symmetrised splitting function, the tail of the analytic calculation moves closer to the Monte Carlo results.

4.4 Comparison to fragmentation functions

The study of heavy-flavour production at high energies is a multi-scale problem and, as already pointed out, logarithms of the ratio of the quark mass to the hard scale of the process can spoil the convergence of the perturbative expansion. Because these logarithmic corrections are related to collinear dynamics, heavy-flavour production cross-sections obey a factorisation theorem and they can be written as the convolution of process-dependent partonic (massless) coefficient functions with universal heavy-flavour fragmentation functions. Fragmentation functions obey DGLAP evolution equations with timelike splitting functions, which allow one to resum these large logarithmic corrections to all perturbative orders.

The initial condition for heavy quark fragmentation functions can be computed in perturbation theory, as originally pointed out in Refs. [27,28], where the next-to-leading order (NLO) computation in QCD was presented. The NNLO corrections were computed later in Refs. [29,30]. Furthermore, the initial condition of the evolution is affected by soft logarithms, that can be resummed to all orders too [31–34,36]. Fragmentation functions for b (or c) quarks are usually then supplemented with non-perturbative corrections before being compared to experimental data. In the context of heavy-flavour production in e^+e^- collision, one of the most widely studied observables is perhaps the energy fraction x of the heavy quark (or hadron) with respect to the energy of the incoming beam, in the centre-of-mass frame, see [84] (and references therein) for a recent review.

Collinear factorisation ensures that fragmentation functions are universal objects and so they can be used, in principle, to describe heavy flavours in hadronic collisions. However, a few changes in their definition are usually required

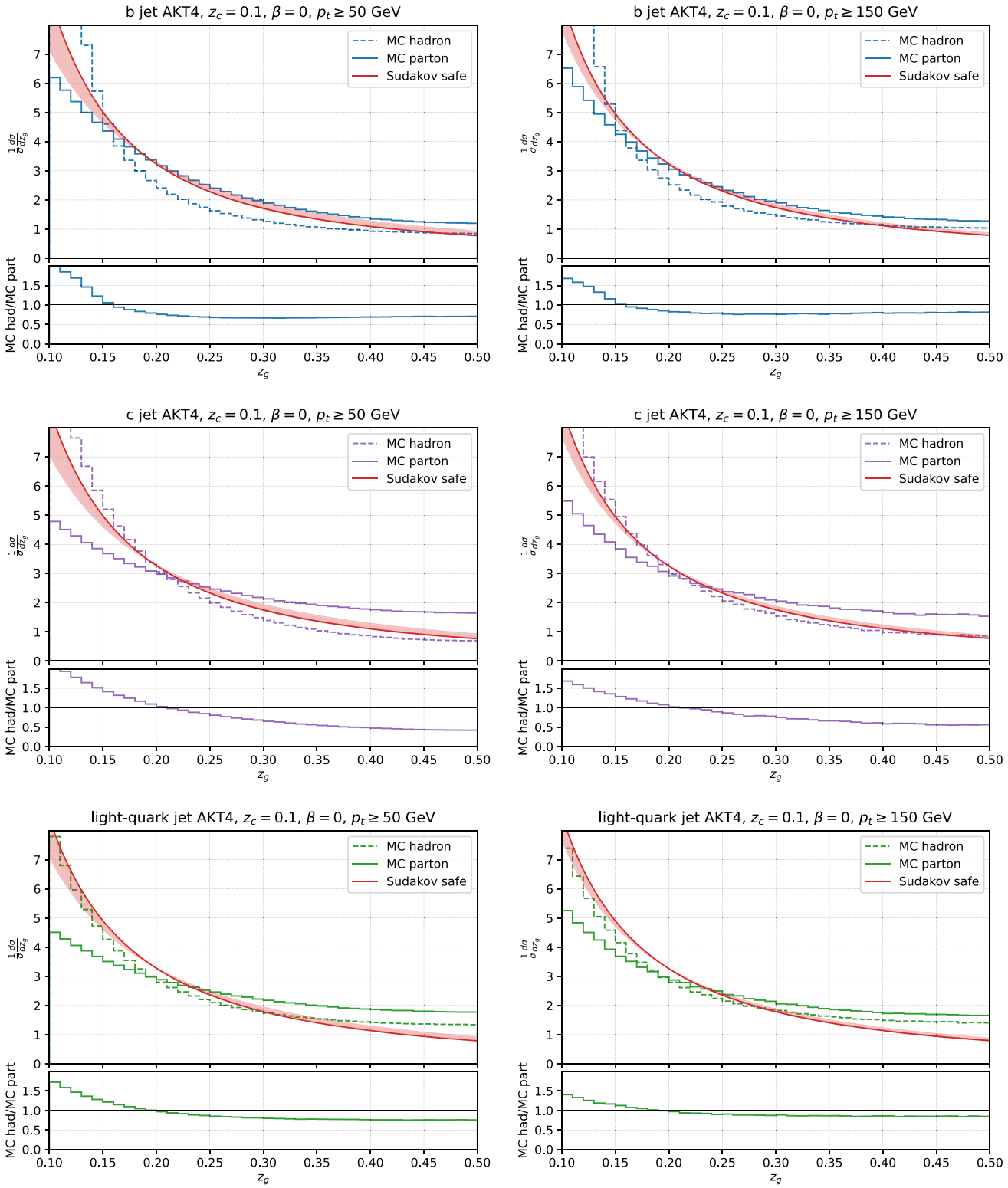

Fig. 4 The groomed momentum fraction (z_g) distribution for b-jets (top), c-jets (middle) and light-quark jets (bottom). In all plots, we show our resummed result, as well as the one obtained with the Monte Carlo event generator HERWIG, both at parton and hadron level. Jets are selected with the anti-k_t algorithm with $R_0 = 0.4$, and groomed with Soft Drop with $z_c = 0.1$ and $\beta = 0$. The plots on the left are for $p_t \geq 50$ GeV, while the ones on the right are for $p_t \geq 150$ GeV. The uncertainty bands for the analytic predictions are obtained by varying the resummation scale by a factor of two above and below the hard scale $p_t R_0$, i.e. $\mu_R \in \left[\frac{p_t R_0}{2}, 2 p_t R_0\right]$

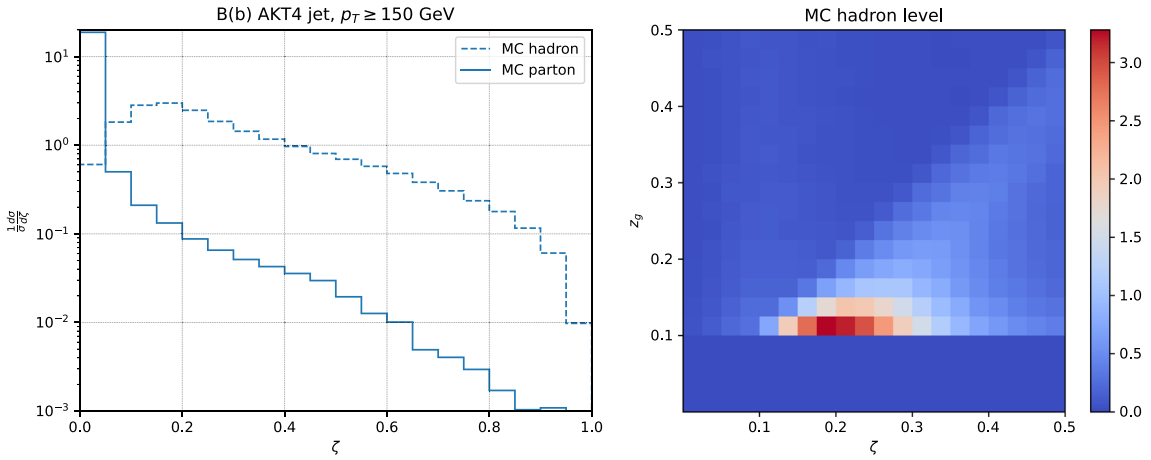

Fig. 5 Monte Carlo studies of the fragmentation function variable x, obtained with HERWIG. The plot on the left shows the $\zeta = 1 - x$ distribution for b-jets, at parton- and hadron-level. The plot on the right shows the correlation between z_g (with $z_c = 0.1$, and $\beta = 0$) and ζ, at hadron level. In both cases, we have $p_t \geq 150$ GeV, and $R_0 = 0.4$

from a practical point of view. For instance, although is possible to study directly the properties of identified B or D decays, the LHC experiments typically measure properties of the B and D hadrons within jets [85–88]. Theoretical studies of heavy quarks fragmenting in jets have been performed using effective field theories [89–92].

In hadronic collisions, one usually measures projections of the hadron momentum with respect to the jet one (see e.g. [85]) or the transverse momentum fraction of the heavy flavour with respect to the jet transverse momentum (see e.g. [86]):

$$x = \frac{p_{ti}}{p_t}, \quad i = b, c. \tag{47}$$

Because $x = 1$ at the Born level, if we want to study soft and collinear emissions, it is useful to define $\zeta = 1 - x$.

In what follows, we would like to study the correlation, if any, between the variable ζ, which measures the departure from Born kinematics of the identified heavy flavour within a jet, to z_g (with $\beta = 0$), which probes the kinematics of the splitting that passes Soft Drop, in a heavy-flavour tagged jet.

Let us start with the $\mathcal{O}(\alpha_S)$ calculation, in the quasi-collinear limit. The result for z_g can be immediately read off Eq. (39) by setting $\beta = 0$:

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.o.})}}{\mathrm{d}z_g} = \frac{\alpha_S}{2\pi}\int_{\theta_i^2}^1 \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}\,\mathscr{P}_{gi}\left(z_g, \bar{\theta}^2\right)\Theta(z_g - z_c). \tag{48}$$

The computation of the momentum fraction ζ, in the same approximation, is rather straightforward. We consider the emission of a gluon with momentum fraction z in the quasi-collinear limit. If the gluon is emitted within the jet, then $z = \zeta$, i.e. the $b(c)$ quark has momentum fraction $1 - z = x$. If the gluon is emitted outside the jet, it does not contribute to the observable and x remains equal to unity.[3] We have

$$\begin{aligned}\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma_i^{(\mathrm{f.o.})}}{\mathrm{d}\zeta} &= \frac{\alpha_S}{2\pi}\int_0^1 \frac{\mathrm{d}\theta^2}{\theta^2 + \theta_i^2}\int_0^1 \mathrm{d}z\, P_{gi}\left(z, (z\theta p_t R_0)^2\right) \\ &\quad \delta(z - \zeta) \\ &= \frac{\alpha_S}{2\pi}\int_{\theta_i^2}^1 \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}\,\mathscr{P}_{gi}\left(\zeta, \bar{\theta}^2\right),\end{aligned} \tag{49}$$

where the second equation holds up to power corrections in the mass. Eq. (49) coincides with Eq. (48), for $\zeta > z_c$, i.e. $x < 1 - z_c$. We conclude that $1 - x$ and z_g are fully correlated at this perturbative order. It is interesting to investigate whether this remains true beyond that. We study this problem by performing a Monte Carlo study.

Before doing so, we study the $\zeta = 1 - x$ spectrum using the same HERWIG simulation setup described for our previous analyses. The results are shown in Fig. 5, on the left, where we plot the ζ distribution at parton- and hadron-level, for b-jets with $p_t \geq 150$ GeV. We note that non-perturbative effects are large. This is interesting because, while fragmentation functions are under better perturbative control than Soft Drop observables, the latter ones seem to be more robust against non-perturbative corrections.

Finally, the right-hand plot of Fig. 5 shows the two dimensional distribution for z_g (with $z_c = 0.1$ and $\beta = 0$) and ζ, again for $p_t \geq 150$ GeV, at hadron-level. We note that the $\mathcal{O}(\alpha_S)$ correlation between these two observables is not maintained when higher-order corrections and non-perturbative effects are included. Therefore, we conclude that

[3] This is different in e^+e^- collision where the energy fraction is defined globally and gluon radiation always contributes.

z_g and $\zeta = 1 - x$ can offer different handles to study heavy-flavour dynamics at the LHC.

5 Conclusions and outlook

In this work, we have considered heavy-flavour jets, namely b- and c-jets. In particular, we have studied heavy-flavour jets in conjunction with the Soft Drop grooming algorithm. Exploiting resummed perturbation theory, we have computed the NLL distribution for the groomed jet radius θ_g, for both b- and c-jets. Our calculation accounts for the dead-cone effect due to the heavy-quark mass, as well as m_b and m_c thresholds in the running coupling, which is frozen at a non-perturbative scale $\Lambda = 1$ GeV. We also discuss the role of clustering logarithms, first computed in [44]. The presence of many scales originates many cases and regions to be considered. We have listed all of them and shown numerical results for representative cases. We have compared our findings for the θ_g distribution with simulated data obtained with Monte Carlo event generators, assessing the role of non-perturbative corrections, such as hadronisation and the Underlying Event. We have found good agreement between resummed perturbation theory and Monte Carlo simulation, for the $\beta = 0$ case, even for relatively low values of the hard scale, i.e. $p_t = 50$ GeV and $R_0 = 0.4$. The agreement worsens if positive values of the angular exponent β are considered.

We have also considered the momentum fraction z_g of the first emission that passes Soft Drop. Our calculation generalises to the massive case the conditional-probability approach of [38] and it allows for a resummation of mass effects. While our results hold for any $\beta \geq 0$, we have focussed our numerical investigations on the $\beta = 0$ case. This choice is motivated by a few reasons. First, this is the value for which, at least to lowest order, one has a clear factorisation of the conditional probability expression, leading to a distribution proportional to the splitting function. For $\beta > 0$, the dependence on the splitting function is smeared out. Second, it is for the $\beta = 0$ case that one expects to find more similarities to the fragmentation variable x, and third, from the study of the θ_g distribution, we have found that the $\beta = 0$ is under better theoretical control and more sensitive to dead cone effects. We have compared our numerical results to Monte Carlo simulations. As for the massless case, the result is driven by the QCD splitting function and it is largely insensitive to non-perturbative effects.

Finally, we have compared the z_g distribution to the fragmentation variable x, which measures the ratio of the heavy quark (hadron) transverse momentum to the jet p_t. We performed the analytic calculation of both distributions at $\mathcal{O}(\alpha_S)$, showing that they lead to the same results for $z_g = 1 - x > z_c$. We have investigated the all-order behaviour of these observables using Monte Carlo simulations and discovered that the parton-shower significantly dilutes this correlation. However, we have found that, in contrast to z_g, the x receives sizeable non-perturbative corrections. In the future, it would be interesting to repeat these studies with the modified version of the declustering procedure [93], whereby one follows the flavoured branch.

In this work, we have performed a detailed theoretical study of the effect of Soft Drop grooming on heavy-flavour jets. Before being able to compare to experimental data, such as the ones collected by the ALICE collaboration [40], a few steps need to be taken. We are going to implement our calculation in the resummation plugin to SHERPA [94,95] in order to match our results to NLO theoretical predictions, with fiducial cuts, supplemented with non-perturbative corrections, as done for instance in [96–101]. In this context, it would be also interesting to lift the small-z_c limit and investigate, in the $\beta = 0$ case, flavour-changing contributions that may induce radiation into the dead-cone region. Finally, it would be important to improve the accuracy of the z_g calculation by including the resummation of z_g and z_c logarithms, as done in [45].

Acknowledgements We thank Oleh Fedkevych, Silvia Ferrario Ravasio, Ezra Lesser, Davide Napoletano, Giovanni Ridolfi, Gregory Soyez, and Maria Ubiali for many inspiring discussions and comments on the manuscript. We acknowledge support from the IPPP DIVA fellowship program and thank the physics department at the University of Cambridge for hospitality during the course of this work. AG and SM would like to thank the Erwin-Schrödinger International Institute for Mathematics and Physics at the University of Vienna for partial support during the Programme "Quantum Field Theory at the Frontiers of the Strong Interactions", July 31–September 1, 2023. SC would like to thank the University of Amsterdam and the Delta-ITP programme for support during the course of this work. We also thank the Galileo Galilei Institute (GGI) for Theoretical Physics for the hospitality and the INFN for partial support during the completion of this work.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: This manuscript has no associated data.]

Funded by SCOAP[3].

Appendix A: Running coupling integrals with quark-mass thresholds

Throughout the paper, we have made use of the strong coupling in the Catani–Marchesini–Webber (CMW) scheme [47]:

$$\alpha_{\rm S}^{\rm CMW}(k_t^2) = \alpha_{\rm S}(k_t^2)\left(1+\alpha_{\rm S}(k_t^2)\frac{K^{(n_f)}}{2\pi}\right), \tag{A.1}$$

where, in turn, $\alpha_{\rm S}(k_t^2)$ is in the decoupling scheme:

$$\begin{aligned}\alpha_{\rm S}(k_t^2) = {}& \alpha_{\rm S}^{(5)}(k_t^2)\Theta(k_t^2-m_b^2)\\ &+\alpha_{\rm S}^{(4)}(k_t^2)\Theta(k_t^2-m_c^2)\Theta(m_b^2-k_t^2)\\ &+\alpha_{\rm S}^{(3)}(k_t^2)\Theta(k_t^2-\Lambda^2)\Theta(m_c^2-k_t^2)\\ &+\alpha_{\rm S}^{(3)}(\Lambda^2)\Theta(\Lambda^2-k_t^2),\end{aligned} \tag{A.2}$$

and

$$K^{(n_f)} = C_{\rm A}\left(\frac{67}{18}-\frac{\pi^2}{6}\right)-\frac{5}{9}n_f. \tag{A.3}$$

In the above equation, $\alpha_{\rm S}^{(n_f)}$ is the two-loop running coupling with n_f active flavours. As a prescription to deal with the non-perturbative region, we have decided to freeze the coupling below the scale $\Lambda \simeq 1$ GeV.

In order to evaluate the integrals over the running coupling, we express each $\alpha_{\rm S}^{(n_f)}(k_t^2)$ appearing in Eq. (A.2) in terms of the value of the strong coupling at the hard scale $p_t R_0$, which we assume to be above the b-quark mass. Requiring continuity at the two quark-mass thresholds, at two loops, we find

$$\alpha_{\rm S}^{(5)}(k_t^2) = \frac{\alpha_{\rm S}}{1+\nu_5}\left(1-\alpha_{\rm S}\frac{\beta_1^{(5)}}{\beta_0^{(5)}}\frac{\log(1+\nu_5)}{1+\nu_5}\right), \tag{A.4}$$

$$\begin{aligned}\alpha_{\rm S}^{(4)}(k_t^2) = {}& \frac{\alpha_{\rm S}}{1+\nu_4-\delta_{54}}\left(1-\alpha_{\rm S}\frac{\beta_1^{(4)}}{\beta_0^{(4)}}\frac{\log(1+\nu_4-\delta_{54})}{1+\nu_4-\delta_{54}}\right)\\ &-\left(\frac{\beta_1^{(5)}}{\beta_0^{(5)}}-\frac{\beta_1^{(4)}}{\beta_0^{(4)}}\right)\log\left(1-\lambda_{\xi_b}^{(5)}\right)\frac{\alpha_{\rm S}^2}{(1+\nu_4-\delta_{54})^2},\end{aligned} \tag{A.5}$$

$$\begin{aligned}\alpha_{\rm S}^{(3)}(k_t^2) = {}& \frac{\alpha_{\rm S}}{1+\nu_3-\delta_{54}-\delta_{43}}\\ &\times\left(1-\alpha_{\rm S}\frac{\beta_1^{(3)}}{\beta_0^{(3)}}\frac{\log(1+\nu_3-\delta_{54}-\delta_{43})}{1+\nu_3-\delta_{54}-\delta_{43}}\right)\\ &-\alpha_{\rm S}^2\left(\frac{\beta_1^{(4)}}{\beta_0^{(4)}}-\frac{\beta_1^{(3)}}{\beta_0^{(3)}}\right)\frac{\log\left(1-\delta_{54}-\lambda_{\xi_c}^{(4)}\right)}{(1+\nu_3-\delta_{54}-\delta_{43})^2}\\ &-\alpha_{\rm S}^2\left(\frac{\beta_1^{(5)}}{\beta_0^{(5)}}-\frac{\beta_1^{(4)}}{\beta_0^{(4)}}\right)\frac{\log\left(1-\lambda_{\xi_b}^{(5)}\right)}{(1+\nu_3-\delta_{54}-\delta_{43})^2},\end{aligned} \tag{A.6}$$

where $\alpha_{\rm S} = \alpha_{\rm S}^{(5)}(p_t^2 R_0^2)$ and we have introduced

$$\nu_{n_f} = \alpha_{\rm S}\beta_0^{(n_f)}\log\frac{k_t^2}{p_t^2R_0^2}, \quad n_f = 3,4,5, \tag{A.7}$$

$$\lambda_{\xi_b}^{(n_f)} = \alpha_{\rm S}\beta_0^{(n_f)}\log\frac{1}{\xi_b}, \quad n_f = 4,5, \tag{A.8}$$

$$\lambda_{\xi_c}^{(n_f)} = \alpha_{\rm S}\beta_0^{(n_f)}\log\frac{1}{\xi_c}, \quad n_f = 3,4, \tag{A.9}$$

$$\delta_{54} = \lambda_{\xi_b}^{(5)}-\lambda_{\xi_b}^{(4)}, \quad \delta_{43} = \lambda_{\xi_c}^{(4)}-\lambda_{\xi_c}^{(3)}. \tag{A.10}$$

The two-loop coefficients of the QCD β-function are

$$\beta_0^{(n_f)} = \frac{11C_{\rm A}-2n_f}{12\pi}, \ \beta_1^{(n_f)} = \frac{17C_{\rm A}^2-5C_{\rm A}n_f-3C_{\rm F}n_f}{24\pi^2}, \tag{A.11}$$

with $C_{\rm A}=3$ and $C_{\rm F}=\frac{4}{3}$.

All the expressions we have to compute can be cast as a double integration over the emission's transverse momentum and its angle with respect to the hard quark, at fixed transverse momentum. Furthermore, it proves convenient to change integration variables to logarithmic ones, namely ν and the emission's rapidity. This way, the integral over rapidity can always be written as $L_0 + c\nu$ with L_0 and c independent of ν. Therefore, all these expressions have the form:

$$\begin{aligned}I^{(n_f)}(\lambda_a,\lambda_b,L_0,c) = {}& \frac{C_{\rm F}}{\left(\alpha_{\rm S}\beta_0^{(n_f)}\right)^2}\int_{-\lambda_b}^{-\lambda_a}{\rm d}\nu\,\frac{\alpha_{\rm S}^{\rm CMW}(\bar{k}_t^2)}{2\pi}\\ &\times(L_0+c\nu),\\ \bar{k}_t = {}& p_t R_0\exp\frac{\nu}{\alpha_{\rm S}\beta_0^{(n_f)}}.\end{aligned} \tag{A.12}$$

The limits of integration, as well as L_0, are linear combinations of $\lambda^{(n_f)} = \alpha_{\rm S}\beta_0^{(n_f)}\log\frac{1}{\vartheta_{g,i}^2}$, $\lambda_{\rm cut}^{(n_f)} = \alpha_{\rm S}\beta_0^{(n_f)}\log\frac{1}{z_c}$, $\lambda^{(\rm NP)} = \alpha_{\rm S}\beta_0^{(3)}\log\frac{(p_tR_0)^2}{\Lambda^2}$ and $\lambda_{\xi_i}^{(n_f)}$, defined above. In contrast, c can only assume the values: $\left\{\frac{1}{1+\beta}, -\frac{\beta}{1+\beta}, -1, 0\right\}$.

As long as $k_t^2 > \Lambda^2$, the integral in Eq. (A.12) can be written as:

$$\begin{aligned}I^{(n_f)}(\lambda_a,\lambda_b,L_0,c) = {}& \frac{C_{\rm F}}{2\pi\beta_0^{(n_f)2}}\left(\frac{1}{\alpha_{\rm S}}I_1(\lambda_a,\lambda_b,L_0,c)\right.\\ &\left.+\frac{K_{\rm eff}^{(n_f)}}{2\pi}I_2(\lambda_a,\lambda_b,L_0,c)-\frac{\beta_1^{(n_f)}}{\beta_0^{(n_f)}}I_3(\lambda_a,\lambda_b,L_0,c)\right),\end{aligned} \tag{A.13}$$

with:

$$
\begin{aligned}
I_1 &= \left(L_0 - c(1-x^{(n_f)})\right)\log\left(\frac{1-x^{(n_f)}-\lambda_a}{1-x^{(n_f)}-\lambda_b}\right)\\
&\quad -c\,(\lambda_a-\lambda_b)\,,\\
I_2 &= (c(1-x^{(n_f)})-L_0)\frac{\lambda_a-\lambda_b}{(1-x^{(n_f)}-\lambda_a)(1-x^{(n_f)}-\lambda_b)}\\
&\quad +c\log\left(\frac{1-x^{(n_f)}-\lambda_a}{1-x^{(n_f)}-\lambda_b}\right)\\
I_3 &= (c(1-x^{(n_f)})-L_0)\\
&\quad \times\left(\frac{1+\log\left(1-x^{(n_f)}-\lambda_a\right)}{1-x^{(n_f)}-\lambda_a}-\frac{1+\log\left(1-x^{(n_f)}-\lambda_b\right)}{1-x^{(n_f)}-\lambda_b}\right)\\
&\quad +c\left(\frac{1}{2}\log^2(1-x^{(n_f)}-\lambda_a)-\frac{1}{2}\log^2(1-x^{(n_f)}-\lambda_b)\right),
\end{aligned}
\tag{A.14}
$$

and

$$
x^{(n_f)} = \begin{cases} 0 & \text{for } n_f = 5,\\ \delta_{54} & \text{for } n_f = 4,\\ \delta_{54}+\delta_{43} & \text{for } n_f = 3. \end{cases}
\tag{A.15}
$$

The coefficient $K_{\text{eff}}^{(n_f)}$ is defined as:

$$
K_{\text{eff}}^{(n_f)} = \begin{cases} K^{(5)} \quad \text{for } n_f = 5,\\ K^{(4)} - 2\pi\left(\frac{\beta_1^{(5)}}{\beta_0^{(5)}}-\frac{\beta_1^{(4)}}{\beta_0^{(4)}}\right)\log\left(1-\lambda_{\xi_b}^{(5)}\right) \quad \text{for } n_f = 4,\\ K^{(3)} - 2\pi\left(\frac{\beta_1^{(4)}}{\beta_0^{(4)}}-\frac{\beta_1^{(3)}}{\beta_0^{(3)}}\right)\log\left(1-\delta_{54}-\lambda_{\xi_c}^{(4)}\right)\\ -2\pi\left(\frac{\beta_1^{(5)}}{\beta_0^{(5)}}-\frac{\beta_1^{(4)}}{\beta_0^{(4)}}\right)\log\left(1-\lambda_{\xi_b}^{(5)}\right) \quad \text{for } n_f = 3. \end{cases}
\tag{A.16}
$$

When k_t becomes smaller than the non-perturbative scale Λ, we freeze the coupling at Λ, see Eq. (A.2). In this case the result of Eq. (A.12) is straightforward, and reads:

$$
\begin{aligned}
I^{(\text{NP})}(\lambda_a,\lambda_b,L_0,c) &= \frac{C_{\text{F}}}{\left(\alpha_{\text{S}}\beta_0^{(n_f)}\right)^2}\frac{\alpha_{\text{S}}^{\text{CMW}}\left(\Lambda^2\right)}{2\pi}\\
&\quad\times\left(\frac{c}{2}(\lambda_a^2-\lambda_b^2)-L_0(\lambda_a-\lambda_b)\right).
\end{aligned}
\tag{A.17}
$$

In order to achieve NLL accuracy, two further contributions must be considered. These are obtained by integrating the less singular components of the splitting function. At this accuracy, we can evaluate the strong coupling at one loop. We must take into account the hard-collinear term of the splitting function:

$$
\begin{aligned}
B^{(n_f)} &= \frac{B_1}{2\pi\beta_0^{(n_f)}\alpha_{\text{S}}}\int_{-\lambda_b}^{-\lambda_a}\mathrm{d}\nu\;\alpha_{\text{S}}^{(n_f)}(\nu)\\
&= \frac{B_1}{2\pi\beta_0^{(n_f)}}\log\left(\frac{1-x^{(n_f)}-\lambda_a}{1-x^{(n_f)}-\lambda_b}\right),
\end{aligned}
\tag{A.18}
$$

and mass-dependent one:

$$
\begin{aligned}
H^{(n_f)} &= \frac{H_1}{2\pi\beta_0^{(n_f)}\alpha_{\text{S}}}\frac{\theta_i^2}{\vartheta_{g,i}^2}\int_{-\lambda_b}^{-\lambda_a}\mathrm{d}\nu\;\alpha_{\text{S}}^{(n_f)}(\nu)\\
&= \frac{H_1\theta_i^2}{2\pi\beta_0^{(n_f)}\vartheta_{g,i}^2}\log\left(\frac{1-x^{(n_f)}-\lambda_a}{1-x^{(n_f)}-\lambda_b}\right),
\end{aligned}
\tag{A.19}
$$

with $B_1 = -\frac{3}{2}C_{\text{F}}$, $H_1 = -C_{\text{F}}$. In the non-perturbative regions, both integrals in Eqs. (A.18, A.19) reduce to the result in Eq. (A.17) with the replacements $\alpha_{\text{S}}^{\text{CMW}} \to \alpha_{\text{S}}$, $L_0 = 1$, and $c = 0$.

Appendix B: Mapping out the different regions

The integrals that are necessary in order to describe heavy-flavour jet distributions are all of the types described in the previous appendix. What makes the calculation cumbersome is the presence of multiple scales, which implies the appearance of many different regions and cases to be considered.

The Soft Drop condition, the heavy-flavour thresholds and the non-perturbative scales determine 7 different values for the dimensionless transverse momentum $\kappa < 1$ that appears on the vertical axis of the Lund plane in Fig. 1:

$$
\kappa^2 = z_c^2,\ \xi_i,\ z_c^2\xi_i^{1+\beta}, \quad \text{with } i = b, c, \Lambda,
\tag{B.20}
$$

Clearly, we have to take into account additional constraints that will reduce the number of possible cases much below 7!. We have

$$
\begin{cases} \xi_b > \xi_c > \xi_\Lambda,\\ z_c^2\xi_b^{1+\beta} > z_c^2\xi_c^{1+\beta} > z_c^2\xi_\Lambda^{1+\beta},\\ \xi_i > z_c^2\xi_i^{1+\beta}. \end{cases}
$$

Note that this implies that $z_c^2\xi_\Lambda^{1+\beta}$ is always the smallest one. We classify the different cases according to the value of z_c.

(a) The case $z_c^2 > \xi_b$ has the following subcases:

1	ξ_c	ξ_Λ	$z_c^2\xi_b^{1+\beta}$	$z_c^2\xi_c^{1+\beta}$
2	ξ_c	$z_c^2\xi_b^{1+\beta}$	ξ_Λ	$z_c^2\xi_c^{1+\beta}$
3	ξ_c	$z_c^2\xi_b^{1+\beta}$	$z_c^2\xi_c^{1+\beta}$	ξ_Λ
4	$z_c^2\xi_b^{1+\beta}$	ξ_c	ξ_Λ	$z_c^2\xi_c^{1+\beta}$
5	$z_c^2\xi_b^{1+\beta}$	ξ_c	$z_c^2\xi_c^{1+\beta}$	ξ_Λ

where it is understood that the values in the table are ordered from big to small.

(b) The case $\xi_b > z_c^2 > \xi_c$ has the same subcases as (a).

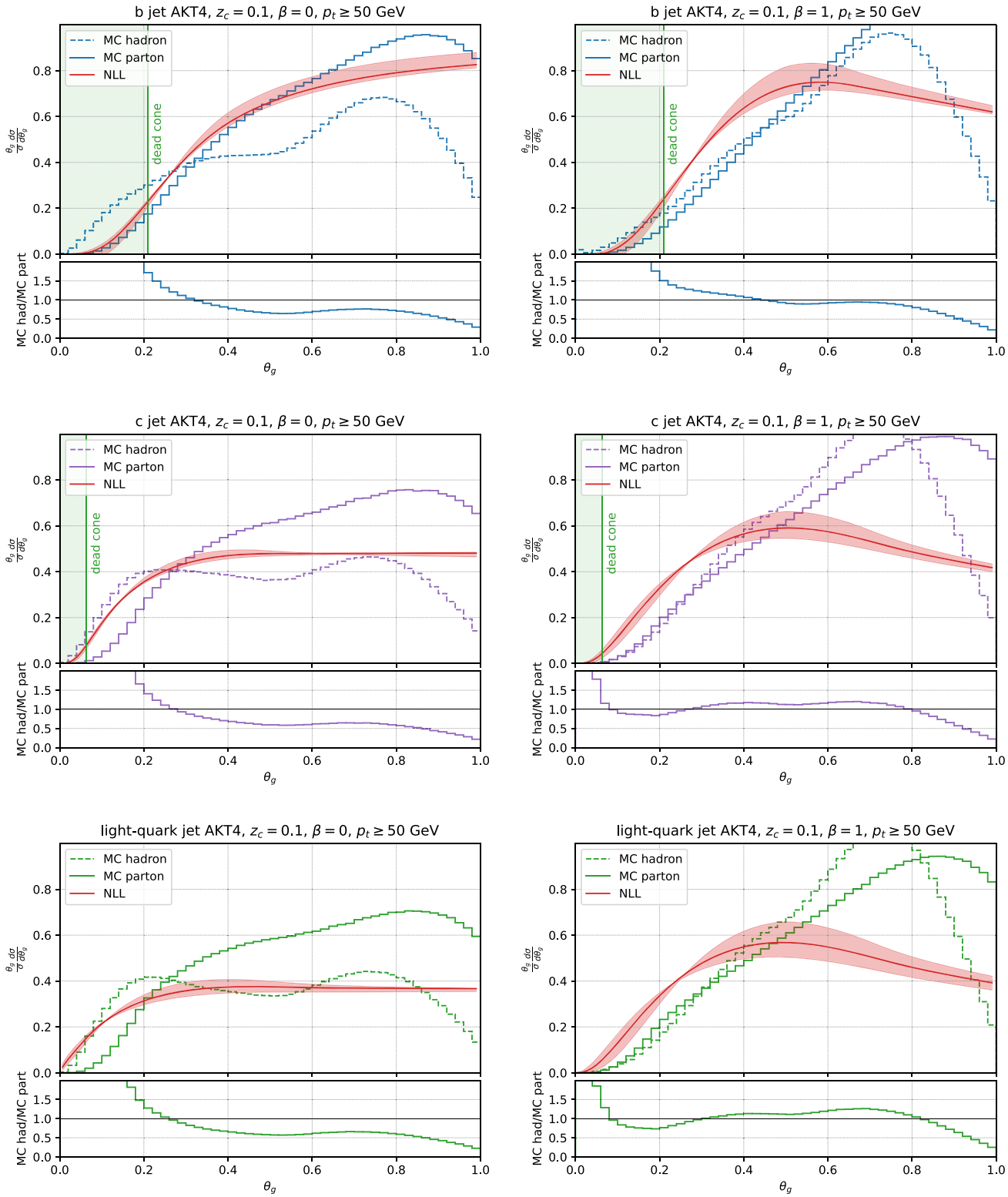

Fig. 6 Same as Fig. 2, but for $p_t \geq 50$ GeV

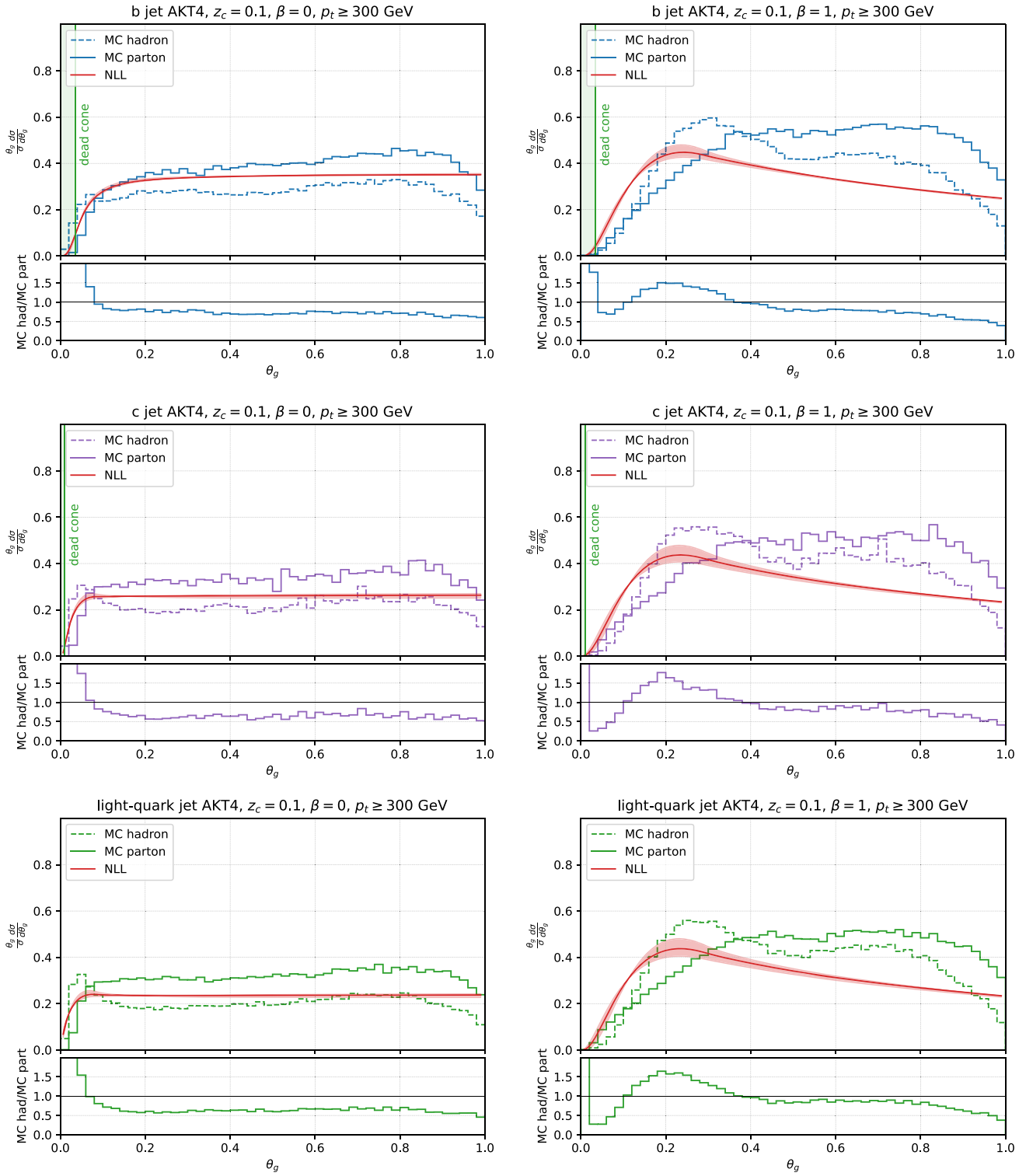

Fig. 7 Same as Fig. 2, but for $p_t \geq 300$ GeV

(c) The case $\xi_c > z_c^2 > \xi_\Lambda$ has 3 subcases:

1	ξ_Λ	$z_c^2\xi_b^{1+\beta}$	$z_c^2\xi_c^{1+\beta}$
2	$z_c^2\xi_b^{1+\beta}$	ξ_Λ	$z_c^2\xi_c^{1+\beta}$
3	$z_c^2\xi_b^{1+\beta}$	$z_c^2\xi_c^{1+\beta}$	ξ_Λ

(d) Finally, for the last case, we can only have $\xi_\Lambda > z_c^2 > z_c^2\xi_b^{1+\beta} > z_c^2\xi_c^{1+\beta}$.

Thus, we have to consider 14 cases. This is what actually happens for light jets and c-jets. For b-jets, the situation is slightly simpler because the dead cone at $\theta_i^2 = \xi_b$ implies that we are not sensitive to the scale $z_c^2\xi_c^{1+\beta}$. Thus, we have only 3 distinct subcases for (a) and (b), 2 for (c) and still 1 for (d), totalling 9 cases for b-jets.

In this work, we have fixed the value of $z_c = 0.1$ and considered only two distinct values for the angular exponent $\beta = 0, 1$. We have also worked with just one value for the anti-k_t jet radius, $R_0 = 0.4$, and fixed the non-perturbative scale $\Lambda = 1$ GeV. This results in only two different hierarchies to be considered. In fact, for $p_t = 150, 300$ GeV we have:

$$z_c^2 > \xi_b > \xi_c > \xi_\Lambda > z_c^2\xi_b^{1+\beta} > z_c^2\xi_c^{1+\beta} > z_c^2\xi_\Lambda^{1+\beta}, \quad \text{(B.21)}$$

while, for $p_t = 50$ GeV, we find

$$\xi_b > z_c^2 > \xi_c > \xi_\Lambda > z_c^2\xi_b^{1+\beta} > z_c^2\xi_c^{1+\beta} > z_c^2\xi_\Lambda^{1+\beta}. \quad \text{(B.22)}$$

Appendix C: Results for the θ_g distribution at low and high transverse momentum

We collect here results for the θ_g distribution at NLL, for two different values of the transverse momentum cut of the jet. In Fig. 6, we show the case $p_t \geq 50$ GeV, while in Fig. 7 we consider a high-p_t cut, namely $p_t \geq 300$ GeV. Details of the plots are the same as in the main text, Fig. 2.

The low-p_t selection allows us to expose the dead cone. In the case of b-jets, we see that the θ_g distribution (for $\beta = 0$) has clearly a distinct shape compared to the light-quark case. However, for such low scales $p_t R_0 \simeq 20$ GeV, we have to deal with larger non-perturbative effects. This is particularly true for the $\beta = 1$ case. The high-p_t selection is under better perturbative control, but in this phase-space region, mass effects are essentially negligible. This is still interesting because, in this context, heavy-flavour identification provides us with a purified sample of quark jets, i.e. it essentially works as a quark/gluon tagger.

References

1. A. Banfi, G.P. Salam, G. Zanderighi, Infrared safe definition of jet flavor. Eur. Phys. J. C **47**, 113–124 (2006). arXiv:hep-ph/0601139
2. S. Caletti, A.J. Larkoski, S. Marzani, D. Reichelt, Practical jet flavour through NNLO. Eur. Phys. J. C **82**(7), 632 (2022). arXiv:2205.01109
3. M. Czakon, A. Mitov, R. Poncelet, Infrared-safe flavoured anti-k$_T$ jets. JHEP **04**, 138 (2023). arXiv:2205.11879
4. R. Gauld, A. Huss, G. Stagnitto, Flavor identification of reconstructed hadronic jets. Phys. Rev. Lett. **130**(16), 161901 (2023). arXiv:2208.11138
5. F. Caola, R. Grabarczyk, M.L. Hutt, G.P. Salam, L. Scyboz, J. Thaler, Flavored jets with exact anti-kt kinematics and tests of infrared and collinear safety. Phys. Rev. D **108**(9), 094010 (2023). arXiv:2306.07314
6. U. Aglietti, L. Di Giustino, G. Ferrera, L. Trentadue, Resummed mass distribution for jets initiated by massive quarks. Phys. Lett. B **651**, 275–292 (2007). arXiv:hep-ph/0612073
7. U. Aglietti, L. Di Giustino, G. Ferrera, A. Renzaglia, G. Ricciardi, L. Trentadue, Threshold resummation in B → X(c) l nu(l) decays. Phys. Lett. B **653**, 38–52 (2007). arXiv:0707.2010
8. U. Aglietti, L. Di Giustino, G. Ferrera, L. Trentadue, Comment on resummation of mass distribution for jets initiated by massive quarks. Phys. Lett. B **670**, 367–368 (2009). arXiv:0804.3922
9. U.G. Aglietti, G. Ferrera, Improved factorization for threshold resummation in heavy quark to heavy quark decays. Eur. Phys. J. C **83**(4), 335 (2023). arXiv:2211.14397
10. S. Fleming, A.H. Hoang, S. Mantry, I.W. Stewart, Jets from massive unstable particles: top-mass determination. Phys. Rev. D **77**, 074010 (2008). arXiv:hep-ph/0703207
11. S. Fleming, A.H. Hoang, S. Mantry, I.W. Stewart, Top jets in the peak region: factorization analysis with NLL resummation. Phys. Rev. D **77**, 114003 (2008). arXiv:0711.2079
12. B. Bachu, A.H. Hoang, V. Mateu, A. Pathak, I.W. Stewart, Boosted top quarks in the peak region with NL3L resummation. Phys. Rev. D **104**(1), 014026 (2021). arXiv:2012.12304
13. A. Jain, I. Scimemi, I.W. Stewart, Two-loop jet-function and jet-mass for top quarks. Phys. Rev. D **77**, 094008 (2008). arXiv:0801.0743
14. A.H. Hoang, C. Lepenik, M. Stahlhofen, Two-loop massive quark jet functions in SCET. JHEP **08**, 112 (2019). arXiv:1904.12839
15. A. Bris, V. Mateu, M. Preisser, Massive event-shape distributions at N^{2}LL. JHEP **09**, 132 (2020). arXiv:2006.06383
16. F. Maltoni, M. Selvaggi, J. Thaler, Exposing the dead cone effect with jet substructure techniques. Phys. Rev. D **94**(5), 054015 (2016). arXiv:1606.03449
17. C. Lee, P. Shrivastava, V. Vaidya, Predictions for energy correlators probing substructure of groomed heavy quark jets. JHEP **09**, 045 (2019). arXiv:1901.09095
18. J. Llorente, J. Cantero, Determination of the b-quark mass m_b from the angular screening effects in the ATLAS b-jet shape data. Nucl. Phys. B **889**, 401–418 (2014). arXiv:1407.8001
19. H.T. Li, I. Vitev, Inverting the mass hierarchy of jet quenching effects with prompt b-jet substructure. Phys. Lett. B **793**, 259–264 (2019). arXiv:1801.00008
20. H.T. Li, Z.L. Liu, I. Vitev, Heavy flavor jet production and substructure in electron-nucleus collisions. Phys. Lett. B **827**, 137007 (2022). arXiv:2108.07809
21. E. Craft, K. Lee, B. Meçaj, I. Moult, Beautiful and charming energy correlators. arXiv:2210.09311
22. L. Cunqueiro, D. Napoletano, A. Soto-Ontoso, Dead-cone searches in heavy-ion collisions using the jet tree. Phys. Rev. D **107**(9), 094008 (2023). arXiv:2211.11789

23. O. Fedkevych, C.K. Khosa, S. Marzani, F. Sforza, Identification of b jets using QCD-inspired observables. Phys. Rev. D **107**(3), 034032 (2023). arXiv:2202.05082
24. Y.L. Dokshitzer, V.A. Khoze, S.I. Troian, On specific QCD properties of heavy quark fragmentation ('dead cone'). J. Phys. G **17**, 1602–1604 (1991)
25. Y.L. Dokshitzer, V.A. Khoze, S.I. Troian, Specific features of heavy quark production. LPHD approach to heavy particle spectra. Phys. Rev. D **53**, 89–119 (1996). arXiv:hep-ph/9506425
26. ALICE Collaboration, S. Acharya et al., Direct observation of the dead-cone effect in quantum chromodynamics. Nature **605**(7910), 440–446 (2022). arXiv:2106.05713. [Erratum: Nature **607**, E22 (2022)]
27. B. Mele, P. Nason, The fragmentation function for heavy quarks in QCD. Nucl. Phys. B **361**, 626–644 (1991). [Erratum: Nucl. Phys. B **921**, 841–842 (2017)]
28. B. Mele, P. Nason, Next-to-leading QCD calculation of the heavy quark fragmentation function. Phys. Lett. B **245**, 635–639 (1990)
29. K. Melnikov, A. Mitov, Perturbative heavy quark fragmentation function through $\mathcal{O}(\alpha_s^2)$. Phys. Rev. D **70**, 034027 (2004). arXiv:hep-ph/0404143
30. A. Mitov, Perturbative heavy quark fragmentation function through $\mathcal{O}(\alpha_s^2)$: gluon initiated contribution. Phys. Rev. D **71**, 054021 (2005). arXiv:hep-ph/0410205
31. M. Cacciari, S. Catani, Soft gluon resummation for the fragmentation of light and heavy quarks at large x. Nucl. Phys. B **617**, 253–290 (2001). arXiv:hep-ph/0107138
32. M. Fickinger, S. Fleming, C. Kim, E. Mereghetti, Effective field theory approach to heavy quark fragmentation. JHEP **11**, 095 (2016). arXiv:1606.07737
33. F. Maltoni, G. Ridolfi, M. Ubiali, M. Zaro, Resummation effects in the bottom-quark fragmentation function. JHEP **10**, 027 (2022). arXiv:2207.10038
34. M. Czakon, T. Generet, A. Mitov, R. Poncelet, NNLO B-fragmentation fits and their application to $t\bar{t}$ production and decay at the LHC. JHEP **03**, 251 (2023). arXiv:2210.06078
35. D. Gaggero, A. Ghira, S. Marzani, G. Ridolfi, Soft logarithms in processes with heavy quarks. JHEP **09**, 058 (2022). arXiv:2207.13567
36. A. Ghira, S. Marzani, G. Ridolfi, A consistent resummation of mass and soft logarithms in processes with heavy flavours. JHEP **11**, 120 (2023). arXiv:2309.06139
37. A.J. Larkoski, S. Marzani, G. Soyez, J. Thaler, Soft drop. JHEP **05**, 146 (2014). arXiv:1402.2657
38. A.J. Larkoski, S. Marzani, J. Thaler, Sudakov safety in perturbative QCD. Phys. Rev. D **91**(11), 111501 (2015). arXiv:1502.01719
39. P. Ilten, N.L. Rodd, J. Thaler, M. Williams, Disentangling heavy flavor at colliders. Phys. Rev. D **96**(5), 054019 (2017). arXiv:1702.02947
40. ALICE Collaboration, S. Acharya et al., Measurements of groomed-jet substructure of charm jets tagged by D0 mesons in proton–proton collisions at s = 13 TeV. Phys. Rev. Lett. **131**(19), 192301 (2023). arXiv:2208.04857
41. M. Cacciari, G.P. Salam, G. Soyez, The anti-k_t jet clustering algorithm. JHEP **04**, 063 (2008). arXiv:0802.1189
42. M. Wobisch, T. Wengler, Hadronization corrections to jet cross-sections in deep inelastic scattering, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279 (1998). arXiv:hep-ph/9907280
43. Y.L. Dokshitzer, G.D. Leder, S. Moretti, B.R. Webber, Better jet clustering algorithms. JHEP **08**, 001 (1997). arXiv:hep-ph/9707323
44. Z.-B. Kang, K. Lee, X. Liu, D. Neill, F. Ringer, The soft drop groomed jet radius at NLL. JHEP **02**, 054 (2020). arXiv:1908.01783
45. P. Cal, K. Lee, F. Ringer, W.J. Waalewijn, The soft drop momentum sharing fraction zg beyond leading-logarithmic accuracy. Phys. Lett. B **833**, 137390 (2022). arXiv:2106.04589
46. B. Andersson, G. Gustafson, L. Lonnblad, U. Pettersson, Coherence effects in deep inelastic scattering. Z. Phys. C **43**, 625 (1989)
47. S. Catani, B.R. Webber, G. Marchesini, QCD coherent branching and semi-inclusive processes at large x. Nucl. Phys. B **349**, 635–654 (1991)
48. M. Dasgupta, A. Fregoso, S. Marzani, G.P. Salam, Towards an understanding of jet substructure. JHEP **09**, 029 (2013). arXiv:1307.0007
49. S. Marzani, L. Schunk, G. Soyez, A study of jet mass distributions with grooming. JHEP **07**, 132 (2017). arXiv:1704.02210
50. A. Banfi, M. Dasgupta, Problems in resumming interjet energy flows with k_t clustering. Phys. Lett. B **628**, 49–56 (2005). arXiv:hep-ph/0508159
51. Y. Delenda, R. Appleby, M. Dasgupta, A. Banfi, On QCD resummation with k(t) clustering. JHEP **12**, 044 (2006). arXiv:hep-ph/0610242
52. Y. Delenda, K. Khelifa-Kerfa, On the resummation of clustering logarithms for non-global observables. JHEP **09**, 109 (2012). arXiv:1207.4528
53. A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, Non-global logarithms and jet algorithms in high-pT jet shapes. JHEP **08**, 064 (2010). arXiv:1004.3483
54. M. Dasgupta, G.P. Salam, Resummation of nonglobal QCD observables. Phys. Lett. B **512**, 323–330 (2001). arXiv:hep-ph/0104277
55. R.B. Appleby, M.H. Seymour, Nonglobal logarithms in interjet energy flow with kt clustering requirement. JHEP **12**, 063 (2002). arXiv:hep-ph/0211426
56. S. Catani, M. Ciafaloni, Many gluon correlations and the quark form-factor in QCD. Nucl. Phys. B **236**, 61–89 (1984)
57. Y.L. Dokshitzer, G. Marchesini, G. Oriani, Measuring color flows in hard processes: beyond leading order. Nucl. Phys. B **387**, 675–714 (1992)
58. S. Catani, S. Dittmaier, Z. Trocsanyi, One loop singular behavior of QCD and SUSY QCD amplitudes with massive partons. Phys. Lett. B **500**, 149–160 (2001). arXiv:hep-ph/0011222
59. S. Catani, S. Dittmaier, M.H. Seymour, Z. Trocsanyi, The dipole formalism for next-to-leading order QCD calculations with massive partons. Nucl. Phys. B **627**, 189–265 (2002). arXiv:hep-ph/0201036
60. M. Czakon, Double-real radiation in hadronic top quark pair production as a proof of a certain concept. Nucl. Phys. B **849**, 250–295 (2011). arXiv:1101.0642
61. M. Czakon, D. Heymes, Four-dimensional formulation of the sector-improved residue subtraction scheme. Nucl. Phys. B **890**, 152–227 (2014). arXiv:1408.2500
62. J. Bellm et al., Herwig 7.2 release note. Eur. Phys. J. C **80**(5), 452 (2020). arXiv:1912.06509
63. A.H. Hoang, S. Plätzer, D. Samitz, On the cutoff dependence of the quark mass parameter in angular ordered parton showers. JHEP **10**, 200 (2018). arXiv:1807.06617
64. S. Platzer, S. Gieseke, Dipole showers and automated NLO matching in Herwig++. Eur. Phys. J. C **72**, 2187 (2012). arXiv:1109.6256
65. C. Bierlich et al., A comprehensive guide to the physics and usage of PYTHIA 8.3. SciPost Phys. Codebases **2022**, 8 (2022). arXiv:2203.11601
66. S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C.P. Yuan, New parton distribution functions from a global analysis of quantum chromodynamics. Phys. Rev. D **93**(3), 033006 (2016). arXiv:1506.07443

67. A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lonnblad, J. Monk, H. Schulz, F. Siegert, Rivet user manual. Comput. Phys. Commun. **184**, 2803–2819 (2013). arXiv:1003.0694
68. M. Cacciari, G.P. Salam, G. Soyez, FastJet user manual. Eur. Phys. J. C **72**, 1896 (2012). arXiv:1111.6097
69. A.J. Larkoski, J. Thaler, Unsafe but calculable: ratios of angularities in perturbative QCD. JHEP **09**, 137 (2013). arXiv:1307.1699
70. P.T. Komiske, E.M. Metodiev, J. Thaler, The hidden geometry of particle collisions. JHEP **07**, 006 (2020). arXiv:2004.04159
71. A. Larkoski, S. Marzani, J. Thaler, A. Tripathee, W. Xue, Exposing the QCD splitting function with CMS open data. Phys. Rev. Lett. **119**(13), 132003 (2017). arXiv:1704.05066
72. A. Tripathee, W. Xue, A. Larkoski, S. Marzani, J. Thaler, Jet substructure studies with CMS open data. Phys. Rev. D **96**(7), 074003 (2017). arXiv:1704.05842
73. Y. Mehtar-Tani, K. Tywoniuk, Groomed jets in heavy-ion collisions: sensitivity to medium-induced bremsstrahlung. JHEP **04**, 125 (2017). arXiv:1610.08930
74. R. Kunnawalkam Elayavalli, K.C. Zapp, Medium response in JEWEL and its impact on jet shape observables in heavy ion collisions. JHEP **07**, 141 (2017). arXiv:1707.01539
75. N.-B. Chang, S. Cao, G.-Y. Qin, Probing medium-induced jet splitting and energy loss in heavy-ion collisions. Phys. Lett. B **781**, 423–432 (2018). arXiv:1707.03767
76. G. Milhano, U.A. Wiedemann, K.C. Zapp, Sensitivity of jet substructure to jet-induced medium response. Phys. Lett. B **779**, 409–413 (2018). arXiv:1707.04142
77. ATLAS Collaboration, G. Aad et al., Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} =$ 13 TeV. Phys. Rev. D **101**(5), 052007 (2020). arXiv:1912.09837
78. A Large Ion Collider Experiment, ALICE Collaboration, S. Acharya et al., Measurement of the groomed jet radius and momentum splitting fraction in pp and Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Phys. Rev. Lett. **128**(10), 102001 (2022). arXiv:2107.12984
79. ALICE Collaboration, S. Acharya et al., Exploration of jet substructure using iterative declustering in pp and Pb–Pb collisions at LHC energies. Phys. Lett. B **802**, 135227 (2020). arXiv:1905.02512
80. STAR Collaboration, M.S. Abdallah et al., Differential measurements of jet substructure and partonic energy loss in Au+Au collisions at $\sqrt{S_{NN}} = 200$ GeV. Phys. Rev. C **105**(4), 044906 (2022). arXiv:2109.09793
81. CMS Collaboration, A.M. Sirunyan et al., Measurement of the splitting function in pp and Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Phys. Rev. Lett. **120**(14), 142302 (2018). arXiv:1708.09429
82. STAR Collaboration, K. Kauder, Measurement of the shared momentum fraction z_g using jet reconstruction in p+p and Au+Au collisions with STAR. Nucl. Part. Phys. Proc. **289-290**, 137–140 (2017). arXiv:1703.10933
83. STAR Collaboration, K. Kauder, Measurement of the shared momentum fraction z_g using jet reconstruction in p+p and Au+Au collisions with STAR. Nucl. Phys. A **967**, 516–519 (2017). arXiv:1704.03046
84. G. Corcella, Selected results in heavy-quark fragmentation. Universe **8**(9), 490 (2022). arXiv:2206.11518
85. ATLAS Collaboration, G. Aad et al., Measurement of b-quark fragmentation properties in jets using the decay $B^{\pm} \to J/\psi K^{\pm}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP **12**, 131 (2021). arXiv:2108.11650
86. ATLAS Collaboration, G. Aad et al., Measurements of jet observables sensitive to b-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector. Phys. Rev. D **106**(3) 032008 (2022). arXiv:2202.13901
87. LHCb Collaboration, R. Aaij et al., Study of J/ψ production in jets. Phys. Rev. Lett. **118**(19), 192001 (2017). arXiv:1701.05116
88. ALICE Collaboration, S. Acharya et al., Measurement of the production of charm jets tagged with D^0 mesons in pp collisions at $\sqrt{s} = 7$ TeV. JHEP **08**, 133 (2019). arXiv:1905.02510
89. Y.-T. Chien, Z.-B. Kang, F. Ringer, I. Vitev, H. Xing, Jet fragmentation functions in proton–proton collisions using soft-collinear effective theory. JHEP **05**, 125 (2016). arXiv:1512.06851
90. R. Bain, L. Dai, A. Hornig, A.K. Leibovich, Y. Makris, T. Mehen, Analytic and Monte Carlo studies of jets with heavy mesons and quarkonia. JHEP **06**, 121 (2016). arXiv:1603.06981
91. R. Bain, L. Dai, A. Leibovich, Y. Makris, T. Mehen, NRQCD confronts LHCb data on quarkonium production within jets. Phys. Rev. Lett. **119**(3), 032002 (2017). arXiv:1702.05525
92. Z.-B. Kang, J.-W. Qiu, F. Ringer, H. Xing, H. Zhang, J/ψ production and polarization within a jet. Phys. Rev. Lett. **119**(3), 032001 (2017). arXiv:1702.03287
93. L. Cunqueiro, M. Płoskoń, Searching for the dead cone effects with iterative declustering of heavy-flavor jets. Phys. Rev. D **99**(7), 074027 (2019). arXiv:1812.00102
94. E. Gerwick, S. Hoeche, S. Marzani, S. Schumann, Soft evolution of multi-jet final states. JHEP **02**, 106 (2015). arXiv:1411.7325
95. D. Reichelt, Predictions to all orders in perturbative quantum chromodynamics for high energy collider experiments. PhD thesis, Gottingen U., Georg-August-Universität Göttingen, Gottingen U. (2021)
96. S. Marzani, D. Reichelt, S. Schumann, G. Soyez, V. Theeuwes, Fitting the strong coupling constant with soft-drop thrust. JHEP **11**, 179 (2019). arXiv:1906.10504
97. N. Baberuxki, C.T. Preuss, D. Reichelt, S. Schumann, Resummed predictions for jet-resolution scales in multijet production in e^+e^- annihilation. JHEP **04**, 112 (2020). arXiv:1912.09396
98. J. Baron, D. Reichelt, S. Schumann, N. Schwanemann, V. Theeuwes, Soft-drop grooming for hadronic event shapes. JHEP **07**, 142 (2021). arXiv:2012.09574
99. S. Caletti, O. Fedkevych, S. Marzani, D. Reichelt, S. Schumann, G. Soyez, V. Theeuwes, Jet angularities in Z+jet production at the LHC. JHEP **07**, 076 (2021). arXiv:2104.06920
100. D. Reichelt, S. Caletti, O. Fedkevych, S. Marzani, S. Schumann, G. Soyez, Phenomenology of jet angularities at the LHC. JHEP **03**, 131 (2022). arXiv:2112.09545
101. M. Knobbe, D. Reichelt, S. Schumann, (N)NLO+NLL' accurate predictions for plain and groomed 1-jettiness in neutral current DIS. JHEP **09**, 194 (2023). arXiv:2306.17736