

LA-UR-18-26754 (Accepted Manuscript)

## Transverse momentum spectra at threshold for groomed heavy quark jets

Vaidya, Varun P  
Makris, Yiannis

Provided by the author(s) and the Los Alamos National Laboratory (2019-07-12).

**To be published in:** Journal of High Energy Physics

**DOI to publisher's version:** 10.1007/JHEP10(2018)019

**Permalink to record:** <http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-18-26754>

**Disclaimer:**

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

# Transverse momentum spectra at threshold for groomed heavy quark jets

Yiannis Makris and Varun Vaidya

*Theoretical Division, Group T-2, MS B283, Los Alamos National Laboratory,  
P.O. Box 1663, Los Alamos, NM 87545, U.S.A.*

*E-mail:* [yiannis@lanl.gov](mailto:yiannis@lanl.gov), [vvaidya@lanl.gov](mailto:vvaidya@lanl.gov)

**ABSTRACT:** We present the transverse momentum spectrum for a heavy hadron at threshold in a groomed jet initiated by a heavy quark. The cross section is doubly differential in the energy fraction of an identified heavy hadron in the jet and its transverse momentum measured with respect to the groomed (recoil free) jet axis. The grooming is implemented using a soft-drop grooming algorithm and helps us in mitigating the effects of Non-Global logarithms and pile up. For the particular case of a  $B$  meson, we identify two distinct regimes of the transverse momentum spectrum and develop an EFT within the formalisms of Soft Collinear Effective Theory (SCET) and Heavy Quark Effective Theory (HQET) for each of these regions. We show how each region can be matched smoothly into the other to provide a prediction for the perturbative transverse momentum spectrum. The EFT also predicts the scaling behavior of the leading non-perturbative power corrections and implements a simple shape function to account for hadronization. We work in the threshold region where the heavy hadron carries most of the energy of the jet since in this regime, we have a very good discriminating power between heavy quark and gluon initiated jets. We observe that the shape of the spectrum is independent of the energy of the jet over a large range of transverse momentum. We propose that this spectrum can be used as a probe of evolution for heavy quark TMD fragmentation function. At the same time, it can be treated as a jet substructure observable for probing Quark-Gluon Plasma (QGP).

**KEYWORDS:** Jets, QCD Phenomenology

**ARXIV EPRINT:** [1807.09805](https://arxiv.org/abs/1807.09805)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Factorization</b>	<b>3</b>
2.1	Soft-drop grooming algorithm	4
2.2	Momentum modes	4
<b>3</b>	<b>Resummation</b>	<b>7</b>
3.1	Region 2: $q_{\perp} \gg m(1-z)$	7
3.2	Region 1: $q_{\perp} \sim m(1-z)$	9
3.3	The transverse momentum spectrum	10
<b>4</b>	<b>Comparison with simulation</b>	<b>11</b>
<b>5</b>	<b>Non-perturbative corrections</b>	<b>13</b>
<b>6</b>	<b>Conclusion</b>	<b>14</b>
<b>A</b>	<b>One loop results</b>	<b>16</b>
A.1	GSoft	16
A.2	BHQET	16
A.2.1	Region 1: $q_{\perp} \sim m(1-z)$	16
A.2.2	Region 2: $q_{\perp} \gg m(1-z)$	18
A.3	HQETm	19
A.4	CSoft	19
<b>B</b>	<b>Outline of factorization</b>	<b>20</b>

---

**1 Introduction**

The transverse momentum spectrum with respect to a chosen jet axis of an energetic or massive color-singlet state has been widely studied in literature and is frequently used in probing quantum chromodynamics (QCD) as well as the factorization theorems that separate the long distance physics from ultraviolet hard processes. Hadron-hadron or hadron-lepton collisions (that is, Drell-Yan like spectra and semi-inclusive production of a hadron in deep inelastic scattering) are often used as a tool to study the three dimensional structure of partons inside a hadron. The infrared physics is usually encoded into distribution functions sensitive to the produced transverse momentum and the energy of the hard color singlet state [1–3] with appropriate soft factors and subtractions [4]. This extends the traditional factorization of hadronic structure in terms of collinear Parton Distribution

Functions (PDFs) [5–8] to what are referred to as Transverse Momentum Dependent PDF's or TMDPDF's. When the color-singlet state is within a final state jet not aligned with the beam (for example, observing the transverse momentum spectrum of a final state hadron within an  $e^+e^-$ -collision using hemispherical jets [9]), we can likewise extend the notion of collinear Fragmentation Functions (FFs) to include the relative motion of the hadron with respect to all of the other jet constituents in the form of TMDFF's.

Such fragmentation processes are often studied within a *jet* identified with specific jet algorithms such as anti-kT.<sup>1</sup> Since jets are simply a pattern of radiation that is most likely to occur, there is some arbitrariness in the definition of the jet that can be used. While all reasonable jet definitions usually group together the same energetic radiation into the jet region, the soft radiation that is included inside the jet varies between the different jet algorithms used and this can spoil the equivalence of the TMD-evolution between final state (TMD-FF's) and initial state processes (TMD-PDF's).

There are issues related to soft correlations that span the whole event, entangling the pattern of soft radiation within the jet to either the underlying event with multiple parton interactions within the colliding hadrons [27–32], or non-global color correlations arising from out-of-jet radiation radiating back into the measured jet, all of which are color connected back to the hard process [33, 34]. Indeed, both effects could potentially spoil the factorization predictions for TMD-spectra found in [24, 26]. Thus naively, one suspects that only the TMD-evolution of fragmentation processes within hemisphere jets at an  $e^+e^-$  machine could be tied to the TMD-evolution of the  $Z$ -boson spectrum.

Developments in jet substructure<sup>2</sup> have shown that the modified mass drop tagging algorithm (mMDT) or soft-drop grooming procedure robustly removes contamination from both underlying event and non-global color-correlations, see refs. [36–38], and have been applied to study a wide variety of QCD phenomenology within jets [39–47]. Exploiting this fact, we will give a concrete proposal as to how one can observe the TMD spectrum of heavy quarks within these groomed jets, where we specify that we study the transverse momentum of a hadron within the jet with respect to the total momentum of the groomed jet, that is, all particles that pass the mMDT or soft-drop procedure.

Jet substructure observables are playing a significant role in various experiments both at high (LHC) and low energy (e.g. RHIC). The focus is on precision standard model measurements [48–50] as well as searches for new physics [51–54]. At the same time, jet substructure measurements are being used as a probe of the Quark-Gluon Plasma (QGP) medium. QCD jets produced from early stage collisions of beam quarks and gluons from two nuclei play a central role in studying the transport properties of QGP. During their propagation through the hot and dense medium, the interaction between the hard jets and the colored medium will lead to parton energy loss (jet quenching) [55–57]. There have been several experimental signatures of jet energy loss observed at RHIC and the LHC such as modification of reconstructed jets [58–61] and jet substructure [62–64] as compared to the expectations from proton proton ( $pp$ ) collisions. Continued progress relies on achieving

<sup>1</sup>See refs. [10–26] for recent work on fragmentation processes both generating and within jets.

<sup>2</sup>See ref. [35] for a comprehensive review.

a deeper understanding of the dynamics of jets, allowing for subtle features in a jet to be exploited. This understanding has progressed rapidly in recent years, both due to advances in explicit calculations [37, 38, 65–68] as well as due to the development of techniques for understanding dominant properties of substructure observables using analytic [69–71] and machine learning [72–78] approaches.

While the phenomenology of jets initiated by light/massless partons has been studied extensively ([37, 38, 65–68], little attention has been paid to the case of jets initiated by massive quarks. For recent work on jet substructure calculations on top quark jets, see [40, 79] while the issue of heavy quark jet fragmentation has been addressed in [80]. The focus of this paper will be to develop a factorization theorem to study the transverse momentum spectrum of a heavy hadron inside a groomed jet initiated by a heavy quark. We are interested in the regime where the heavy hadron carries most of the energy of the jet. The introduction of mass as a relevant perturbative scale radically changes the structure of radiation and hence the transverse momentum spectrum of a hadron. The form of the factorization now involves a TMD Heavy Quark Fragmentation function (written with the formalism of Heavy Quark Effective Theory (HQET)), whose evolution bears little resemblance to the corresponding massless object. Indeed, one of the most startling predictions we make in this paper is that the shape of transverse momentum spectrum at low (with respect to the hard process)  $q_{\perp}$  is essentially independent of the hard scattering scale.

The outline of this paper is as follows. In section 2 we develop the factorization theorems in two distinct regimes of transverse momenta within the SCET [81–85] and HQET [86] formalisms. The anomalous dimension (given in appendix A) are then used to resum large logarithms in the transverse momentum in section 3. The parton and hadron level spectrum that we obtain is compared against PYTHIA (section 4). We then provide an analysis of the scaling behavior of the leading non-perturbative effects in section 5. We conclude in section 6. The definitions of the operators that appear in our factorization theorem along with the one loop results are provided in appendix A. In appendix B we show some details on the factorization of the cross section.

## 2 Factorization

Consider the simple case of  $e^+e^-$  collisions where the hard annihilation creates a heavy quark, anti-quark pair. The center of mass energy is much larger than the mass of the quark so that we get two highly boosted partons moving back to back. We isolate a hemisphere jet initiated by heavy quark using a suitable jet algorithm. The heavy quark showers and eventually hadronizes. The jet is groomed with a soft-drop grooming algorithm. The heavy hadron is identified and we measure the energy fraction,  $z$ , of the hadron with respect to the jet energy. At the same time, we also measure its transverse momentum ( $q_{\perp}$ ) with respect to the axis of the groomed jet (which is recoil free). We are interested in the regime where the heavy hadron carries most of the jet energy ( $z \sim 0.8$ ).

## 2.1 Soft-drop grooming algorithm

Soft-drop grooming [36] removes contaminating soft radiation from the jet by constructing an angular ordered tree of the jet, and removing the branches at the widest angles which fail an energy requirement. The angular ordering of the jet is constructed through the Cambridge/Aachen (C/A) clustering algorithm [87–91]. As soon as a branch is found that passes the test, it is declared the groomed jet, and all the constituents of the branch are the groomed constituents. At the end of the grooming procedure only the narrow energetic core remains from the original jet. Since at large angles, all collinear energetic radiation is to be found at the center of the jet, no cone is actually imposed to enclose this core. One simply finds the branch whose daughters are sufficiently energetic. Formally the daughters could have any opening angle, though their most likely configuration is collinear.

The strict definition of the algorithm is as follows. Given an ungroomed jet, first we build the clustering history by starting with a list of particles in the jet. At each stage we merge the two particles within the list that are closest in angle.<sup>3</sup> This gives a pseudo-particle, and we remove the two daughters from the current list of particles, replacing them with the merged pseudo-particle. This is repeated until all particles are merged into a single parent. Then we open the tree back up. At each stage of the declustering, we have two branches available, label them  $i$  and  $j$ . We require:

$$\frac{\min\{E_i, E_j\}}{E_i + E_j} > z_{\text{cut}} \left(\frac{\theta_{ij}}{R}\right)^\beta, \tag{2.1}$$

where  $z_{\text{cut}}$  is the modified mass drop parameter,  $\beta$  is the parameter which controls the angularities,  $\theta_{ij}$  is the angle between  $i^{\text{th}}$  and  $j^{\text{th}}$  particle,  $R$  is the jet radius and  $E_i$  is the energy of the branch  $i$ . If the two branches fail this requirement, the softer branch is removed from the jet, and we decluster the harder branch, once again testing eq. (2.1) within the hard branch. The pruning continues until we have a branch that when declustered passes the condition eq. (2.1). All particles contained within this branch whose daughters are sufficiently energetic constitute the groomed jet. Intuitively we have identified the first genuine collinear splitting.

For a hadron-hadron collision, one uses the transverse momentum ( $p_T$ ) with respect to the beam for the condition of eq. (2.1),

$$\frac{\min\{p_{Ti}, p_{Tj}\}}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \left(\frac{\theta_{ij}}{R}\right)^\beta. \tag{2.2}$$

We formally adopt the power counting  $z_{\text{cut}} \ll 1$ , though typically one chooses  $z_{\text{cut}} \sim 0.1$ . See [41] for a study on the magnitude of the power corrections with respect to  $z_{\text{cut}}$  for jet mass distributions. Also for the example we considering here we take  $\beta = 0$ .

## 2.2 Momentum modes

A constrain on the collinearity of the relevant modes is imposed by the fact that we wish to avoid the effects of non-global logs. These will appear if any of our modes contributing

---

<sup>3</sup>This merging is usually taken to be summing the momenta of the particles, though one could use winner-take-all schemes [92–94].

to the measurement become sensitive to the boundary of the jet. Since we assume large jet radius of order 1, the only modes that can be sensitive to the boundary are wide angle modes, i.e., modes for which  $\theta \sim 1$ . Thus to nullify the effect of non-global logs, we need  $q_{\perp}/(E_J(1-z)) \ll 1$ , where  $E_J$  is the energy of the jet, so that all our modes (contributing to the measurement) are collinear and cannot resolve the boundary.

In our current hierarchy, any large angle ( $\theta \sim 1$ ) mode which hopes to pass soft-drop would scale (in light cone co-ordinates) as  $E_J z_{\text{cut}}(1, 1, 1)$ . However since this mode have parametrically large transverse momentum, it must necessarily fail soft-drop. So any correction to the cross section from this region of phase space, will only be a  $z_{\text{cut}}$  (we will assume  $\beta = 0$  for this paper) dependent normalizing factor, but will not influence the shape of the transverse momentum distribution. We left the details of factorization in appendix B and in this section we only show the final result. The cross section can be factorized in the following way:

$$\frac{d\sigma}{d^2\vec{q}_{\perp} dz} = \sigma_0(E_J, z_{\text{cut}}) S_G(E_J z_{\text{cut}}) \times J(q_{\perp}, (1-z), E_J, z_{\text{cut}}, m) \quad (2.3)$$

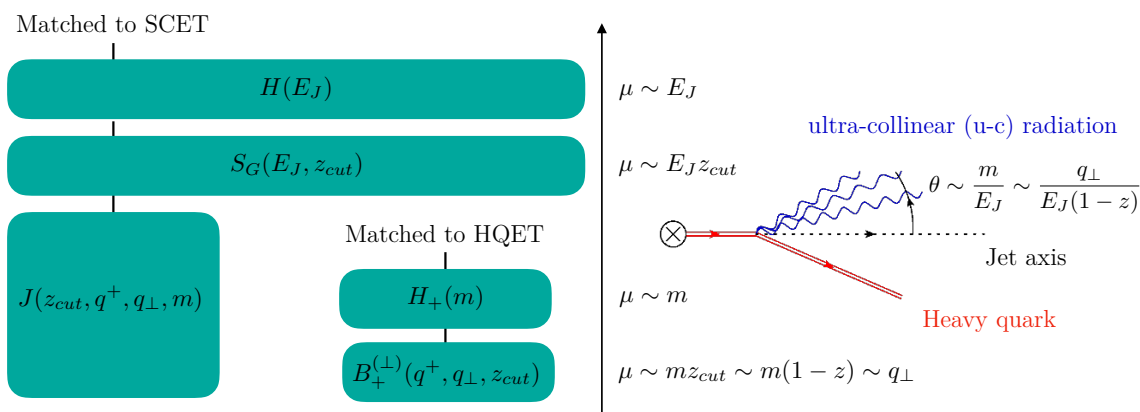
where  $S_G$  is the global soft function defined and evaluated at NLO in A.1, and  $\sigma_0(E_J, z_{\text{cut}})$  is an overall factor describing the hard process and contribution from non-global emissions. The jet function,  $J$ , describes the TMD fragmentation within the hemisphere and is sensitive to the collinear radiation within the hemisphere and contains all the transverse momentum and energy fraction dependence.

All the radiation modes that contribute to the measurement must necessarily pass soft-drop. Hence they must have atleast  $z_{\text{cut}} E_J$  amount of energy. At the same time, all the modes have an upper bound on the energy scaling  $\sim E_J(1-z)$  set by the energy carried by the heavy quark. Since we are working in the regime  $z_{\text{cut}} \sim 1-z$ , this uniquely fixes the energy scaling of *all* modes to be  $z_{\text{cut}} E_J \sim (1-z) E_J$ .

We have one more measurement which is the transverse momentum  $q_{\perp}$  with respect to the groomed jet axis. For a mode to contribute to this measurement, we require that the angle it makes with the jet axis scale as  $\theta \sim q_{\perp}/((1-z)E_J)$ . At the same time, Heavy Quark Effective theory (HQET) tells us that at leading power in the  $1/m$  expansion ( $m$  being the mass of the heavy quark), there is a limit on the angle of any radiation mode with respect to the heavy quark given by  $\theta_{\text{min}} = m/E_J$ . Any radiation off the heavy quark at parametrically lower angles is highly suppressed. We can now consider two cases depending on how  $\theta$  compares with  $\theta_{\text{min}}$ .

- Region 1,  $\theta \sim \theta_{\text{min}}$ : this directly tells us that  $q_{\perp} \sim m(1-z)$  so that transverse momentum  $q_{\perp}$  of the heavy hadron is much smaller than its mass. At the partonic level, we then see that  $q_{\perp} \ll m$ . We then have a single radiation mode that scales as  $p^{\mu} \sim (Q(1-z), q_{\perp}^2/(Q(1-z)), q_{\perp})$  which contributes to both measurements. Equivalently, we can write the scaling of this mode as  $p^{\mu} \sim m(1-z) (Q/m, m/Q, 1)$ , which can be identified as the boosted soft mode of HQET [79] (referred to as the ultra-collinear or u-c mode). For convenience we have defined  $Q = 2E_J$ . The jet function is then matched onto boosted HQET and further factorizes as follows:

$$J(q_{\perp}, (1-z), E_J, z_{\text{cut}}, m) = H(m) \times B_{+}^{(\perp)}(q_{\perp}, E_J(1-z), E_J z_{\text{cut}}, m z_{\text{cut}}) \quad (2.4)$$



**Figure 1.** Hierarchy of scales in region 1 factorization ( $q_\perp \sim m(1-z)$ ). There is only a single radiation mode which contributes to both the  $z$  and  $q_\perp$  measurement which can be identified as the boosted soft mode of HQET (referred to as the ultra-collinear or u-c mode).

where  $B_+^{(\perp)}$  is the boosted HQET jet function (A.2.1) with additional transverse momentum measurement and  $H(m)$  is the matching coefficient from massive SCET to HQET (A.3). The details for the matching procedure are illustrated in appendix B.

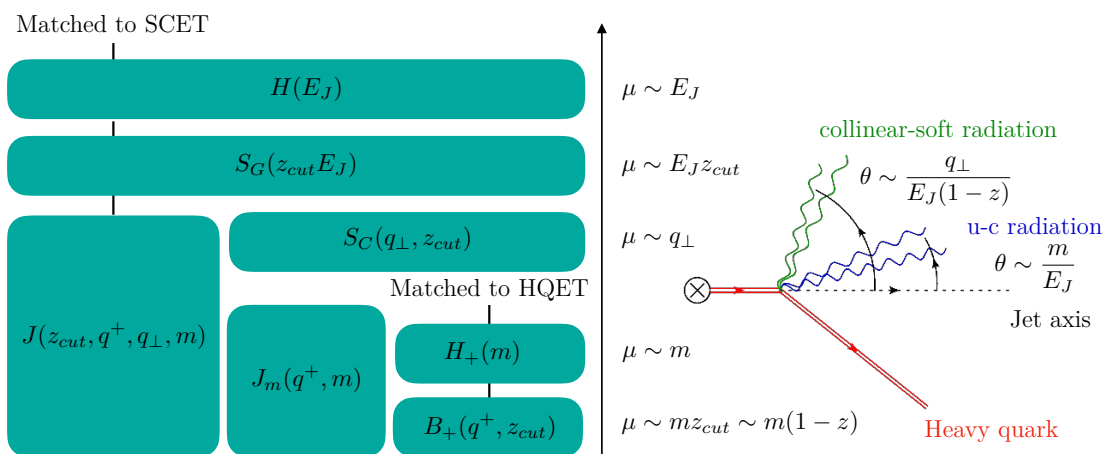
- Region 2,  $\theta \gg \theta_{\min}$ : in this case we define two radiation modes, one that contributes to the transverse momentum of the hadron, and another, which does not. The collinear-soft scales as  $p^\mu \sim (Q(1-z), q_\perp^2/(Q(1-z)), q_\perp)$  and contributes to both measurements. Our condition on  $\theta$  implies that  $q_\perp \gg m(1-z)$ . The other mode which we refer to as the ultra-collinear(u-c), scales as  $p^\mu \sim m(1-z)(Q/m, m/Q, 1)$  and only contributes to the  $z$  measurement. The transverse momentum of this mode  $p_\perp \sim m(1-z)$  is parametrically smaller than  $q_\perp$  and hence does not contribute to the transverse momentum of the heavy hadron. The jet function is then further factorizes as follows:

$$J(q_\perp, (1-z), E_J, z_{\text{cut}}, m) = H(m) \times S_C(E_J z_{\text{cut}}, q^+, q_\perp) \otimes_z B_+(q^+, E_J z_{\text{cut}}, m z_{\text{cut}}) \quad (2.5)$$

where  $B_+$  is the boosted HQET jet function (A.2.2) and  $S_C$  is the collinear soft function (A.4). (For the refactorization of the bHQET jet function see discussion in appendix B.) The convolution  $\otimes_z$  is defined as follows:

$$f(q^+) \otimes_z g(q^+) = \int_0^\infty dq^+ f(q^+) g(2E_J(1-z) - q^+) \quad (2.6)$$

For a jet energy  $\sim 100$  GeV, we are therefore in the range  $q_\perp \sim (0, m_b)$ . However, the range of  $q_\perp$  that can be probed by this EFT can be arbitrarily increased as long as the collinearity condition is maintained. Obviously, for a fixed  $z$ , this can be achieved by increasing the jet energy  $E_J$ . We demonstrate this in figure 5.



**Figure 2.** Hierarchy of scales in region 2 factorization ( $q_\perp \gg m(1-z)$ ). We now have an Collinear-soft mode that contributes to both  $z$  and  $q_\perp$  measurements, while the u-c mode only contributes to the  $z$  measurement.

### 3 Resummation

We have two distinct regimes of  $q_\perp$  and hence two separate factorization theorems. We now proceed to resum large logarithms(if any) in each region and match them to obtain the transverse momentum spectrum over the full range of  $q_\perp$ .

#### 3.1 Region 2: $q_\perp \gg m(1-z)$

In both regions of phase-space we consider here, the jet function involves the HQET hard coefficient  $H_+(m)$ . Although this function does satisfy renormalization group equations (RGEs) we will not consider its the evolution since it only contributes to an overall normalization factor and does not influence the shape of the transverse momentum or the energy fraction spectrum. Therefore here we focus on the RG properties of the collinear-soft function and the HQET jet function which satisfy the following RGE:

$$\frac{d}{d \ln \mu} F(E_J(1-z)) = \int_0^\infty dq^+ \gamma_F(q^+) F(2E_J(1-z) - q^+) \quad (3.1)$$

where  $F$  stands for either  $S_C$  or  $B_+$  and  $\gamma_F$  is the corresponding anomalous dimension. We drop the dependence on any other kinematic variables such as  $m$ ,  $z_{\text{cut}}$  and  $q_\perp$  to improve readability. In momentum space the anomalous dimensions read

$$\begin{aligned} \gamma_{cs}(q^+, E_J z_{\text{cut}}; \mu) &= -2 \frac{\alpha_s(\mu) C_F}{\pi} \frac{\Theta(q^+ - 2E_J z_{\text{cut}})}{q^+} \\ \gamma_B(q^+, E_J z_{\text{cut}}; \mu) &= \frac{\alpha_s C_F}{\pi} \left\{ -2 \ln \left( \frac{\mu}{m z_{\text{cut}}} \right) + 1 + 2 \frac{\theta(q^+ - 2E_J z_{\text{cut}})}{q^+} \right\} \\ &= (\gamma_J - \gamma_+) \delta(q^+) - \gamma_{cs}(q^+, E_J z_{\text{cut}}; \mu) \end{aligned} \quad (3.2)$$

where  $\gamma_J$  is the groomed jet anomalous dimension (see eq. (4.16) in ref. [95]) and  $\gamma_+$  the anomalous dimension of the HQET matching coefficient,  $H_+(m)$  (see eq. (A.19)). We emphasize that the non trivial part of the anomalous dimension for the HQET jet function is the same and opposite sign as the one of the collinear-soft function. This is required from the consistency of anomalous dimensions: since the cross section is independent of the factorization scale, the sum of anomalous dimensions should vanish and therefore the non trivial  $q^+$  terms in  $\gamma_{cs}$  should cancel against the corresponding terms in  $\gamma_B$ . This statement is true for all orders in perturbation theory and here we confirm it at one-loop.

To solve the RG equation, it is easier to work in Laplace space where

$$\frac{d}{d \ln \mu} \tilde{F}(s) = \tilde{\gamma}_F(s) \tilde{F}(s) \quad (3.3)$$

where  $\tilde{F}$  denotes the Laplace transform quantity with respect to  $q^+$  and  $s$  is the Laplace conjugate. The Laplace transform of the collinear-soft anomalous dimension can be written in terms of the incomplete gamma function,

$$\tilde{\gamma}_{cs}(s, E_J z_{\text{cut}}) = -2 \frac{\alpha_s(\mu) C_F}{\pi} \Gamma(0, sQz_{\text{cut}}) \quad (3.4)$$

where  $Q = 2E_J$ . The solution to the RGE is,

$$S_C(s, E_J z_{\text{cut}}, q_\perp; \mu) = \mathcal{U}_{cs}(s; \mu_{sc}, \mu) S_C(s, E_J z_{\text{cut}}, q_\perp; \mu_{sc}) \quad (3.5)$$

where the evolution kernel  $\mathcal{U}_{cs}$  is given by

$$\mathcal{U}_{cs}(s; \mu_{cs}, \mu) = \exp \left[ -2 \frac{C_F}{\pi} \Gamma[0, sQz_{\text{cut}}] \int_{\mu_{cs}}^{\mu} d \ln \mu' \alpha_s(\mu') \right] \quad (3.6)$$

Similarly for the HQET jet function we have

$$B_+(s, E_J z_{\text{cut}}; \mu) = \mathcal{U}_{cs}(s; \mu, \mu_B) \exp \left[ \int_{\mu_B}^{\mu} d \ln \mu' (\gamma_J - \gamma_+) \right] B_+(s, E_J z_{\text{cut}}; \mu_B) \quad (3.7)$$

The scales  $\mu_{cs}$  and  $\mu_B$  are chosen to be the canonical scales for which the logarithms in the fixed order expansion are minimized. As discussed in appendix A those are,

$$\mu_{cs} = q_\perp, \quad \mu_B = m z_{\text{cut}} \quad (3.8)$$

Substituting the fixed order result from eq. (A.23)<sup>4</sup> and keeping only the non-trivial contributions we have for the LL + LO cross section,

$$\frac{d\sigma^{(2)}}{dz d^2 \vec{q}_\perp} = \mathcal{N}(E_J, z_{\text{cut}}, m) \times \mathcal{L}^{-1} \left[ \mathcal{U}_{cs}(s; \mu_{cs}, \mu_B) \frac{\alpha_s C_F}{\pi^2} \frac{1}{q_\perp^2} \Gamma(0, sQz_{\text{cut}}) \right] \quad (3.9)$$

where we absorbed all functions independent of  $s$  and  $q_\perp$  into a single overall normalization factor  $\mathcal{N}$ . Here  $\mathcal{L}^{-1}$  is the inverse Laplace transform with respect to  $s$ . For LL accuracy

---

<sup>4</sup>Note that the HQET jet function at this order only contributes trivial terms proportional to  $\delta^{(2)}(\vec{q}_\perp)$ .

we may keep the first term in the QCD  $\beta$ -function and thus for our final result for the differential cross section in region 2 we get the following,

$$\frac{d\sigma^{(2)}}{dzd^2\vec{q}_\perp} = \mathcal{N}(E_J, z_{\text{cut}}, m) \times \mathcal{L}^{-1} \left[ \frac{1}{q_\perp^2} \Gamma(0, sQz_{\text{cut}}) (r(\mu_{cs}, \mu_B))^{-\frac{4C_F}{\beta_0} \Gamma(0, sQz_{\text{cut}})} \right] \quad (3.10)$$

where  $r(\mu, \mu_0) = \alpha_s(\mu)/\alpha_s(\mu_0) = 1/(1 + \alpha_s(\mu_0)\beta_0/(2\pi) \ln(\mu/\mu_0))$ . Note that  $1/s \sim Q(1-z) \sim Qz_{\text{cut}}$  which prevents us from expanding out the incomplete  $\Gamma$  function in any small parameter. Also, defining  $\tilde{s} = sQ$ , we can write the cross section as

$$\frac{d\sigma^{(2)}}{dzd^2\vec{q}_\perp} = \tilde{\mathcal{N}}(E_J, z_{\text{cut}}, m) \times \mathcal{L}^{-1} \left[ \frac{1}{q_\perp^2} \Gamma(0, \tilde{s}z_{\text{cut}}) (r(\mu_{sc}, \mu_B))^{-\frac{4C_F}{\beta_0} \Gamma(0, \tilde{s}z_{\text{cut}})} \right] \quad (3.11)$$

where  $\tilde{\mathcal{N}} = \mathcal{N}/Q$ . The conjugate variable to  $\tilde{s}$  is just  $(1-z)$ , so that the shape of the  $q_\perp$  distribution is independent of the hard scale  $Q = 2E_J$ . To proceed further, we need to implement the inverse Laplace transform numerically. For future reference, we define the fixed order cross section in this region (with the resummation turned off) as

$$\frac{d\sigma^{(2\text{-FO})}}{dzd^2\vec{q}_\perp} = \tilde{\mathcal{N}}(E_J, z_{\text{cut}}, m) \frac{\alpha_s C_F}{\pi^2} \frac{1}{q_\perp^2} \frac{\theta((1-z) - z_{\text{cut}})}{(1-z)} \quad (3.12)$$

### 3.2 Region 1: $q_\perp \sim m(1-z)$

In this region the collinear-soft and HQET jet function merge into a single function. The corresponding anomalous dimension is given by the sum of  $\gamma_{sc}$  and  $\gamma_B$ . The relevant RGE,

$$\frac{d}{d \ln \mu} B_+^{(\perp)}(\mu) = \gamma_B^{(\perp)} \times B_+^{(\perp)}(\mu) \quad (3.13)$$

where

$$\gamma_B^{(\perp)} = \gamma_J - \gamma_+ = \frac{\alpha_s C_F}{\pi} \left\{ -2 \ln \left( \frac{\mu}{mz_{\text{cut}}} \right) + 1 \right\} \quad (3.14)$$

Since the anomalous dimension does not depend on either of the measured quantities, there are no convolutions involved and thus the solution can be easily written in momentum space,

$$B_+^{(\perp)}(\mu) = \exp \left[ \int_{\mu_B}^{\mu} d \ln \mu' (\gamma_J - \gamma_+) \right] B_+^{(\perp)}(\mu_B) . \quad (3.15)$$

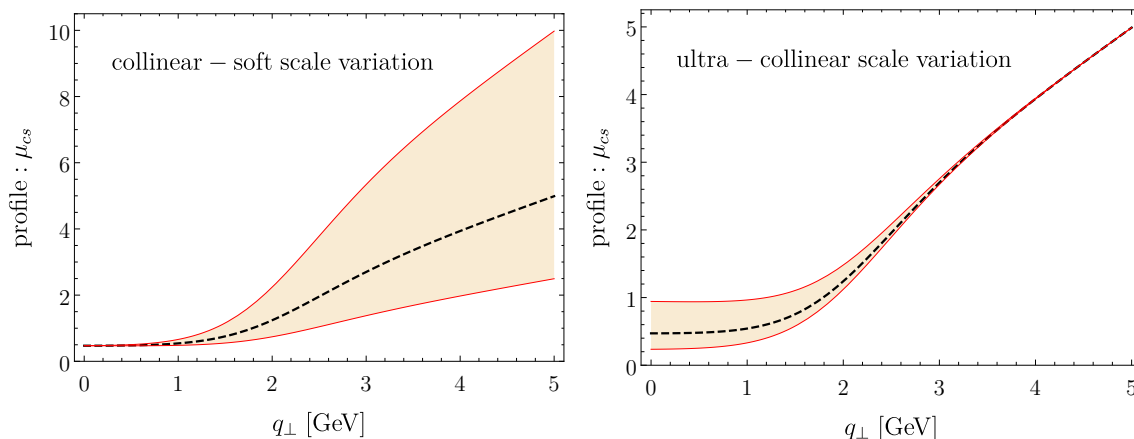
This allow us to write the resummed cross section as follows:

$$\frac{d\sigma^{(1)}}{dzd^2\vec{q}_\perp} = \mathcal{N}(E_J, z_{\text{cut}}, m) \times \left[ \frac{\alpha_s C_F}{\pi^2} \frac{q_\perp^2}{q^+} \frac{\theta(q^+ - Qz_{\text{cut}})}{((q^+)^2 m^2 / Q^2 + q_\perp^2)^2} \right] \quad (3.16)$$

where the normalization factor that appears here,  $\mathcal{N}$ , is the same as in eq. (3.10). Notice that there are no large logarithms (involving the measurement scales) to be resummed so the shape is given by the fixed order result. We can also observe that the cross-section goes to zero smoothly as  $q_\perp$  goes to 0. As in large  $q_\perp$  region, we can express our cross section factoring out the scale  $Q$

$$\frac{d\sigma^{(1)}}{dzd^2\vec{q}_\perp} = \tilde{\mathcal{N}}(E_J, z_{\text{cut}}, m) \times \left[ \frac{\alpha_s C_F}{\pi^2} \frac{q_\perp^2}{(1-z)} \frac{\theta((1-z) - z_{\text{cut}})}{((1-z)^2 m^2 + q_\perp^2)^2} \right] \quad (3.17)$$

So that once again, the shape of the  $q_\perp$  distribution is independent of the hard scale.



**Figure 3.** Profiles in the collinear soft scale  $\mu_S$ , which smoothly turn off the resummation going from region 2 to region 1. The variation in the profiles probes the error band due to missing higher order corrections.

### 3.3 The transverse momentum spectrum

The two regimes of our resummed cross section must smoothly match into the other. It is clear that the only difference between the two regimes is that the  $\ln(mz_{\text{cut}}/q_\perp)$  will be resummed in one ( $q_\perp \gg m(1-z)$ ) and not in the other. On the other hand the power corrections in  $q_\perp/(m/Qq^+)$  are important in the low  $q_\perp$  regimes. So in order to obtain a reliable spectrum in both regions while smoothly interpolating between the two regions we need to turn off the resummation of the  $\ln(mz_{\text{cut}}/q_\perp)$  as we approach intermediate  $q_\perp \sim 2.5$  GeV and match the result to the fixed order (which does not involve any  $q_\perp$  or  $(1-z)$  resummation) cross section. The merging between the two regions can be achieved through multiplicative matching

$$\frac{d\sigma^{(1+2)}}{dzd^2\vec{q}_\perp} = \frac{d\sigma^{(1)}}{dzd^2\vec{q}_\perp} \times \frac{d\sigma^{(2)}}{dzd^2\vec{q}_\perp} / \frac{d\sigma^{(2\text{-FO})}}{dzd^2\vec{q}_\perp} \quad (3.18)$$

where  $d\sigma^{(2\text{-FO})}$  (eq. (3.12)) is the fixed order result in region 2. When the resummation in  $d\sigma^{(2)}$  is turned off at small  $q_\perp$  then is easy to show the following asymptotic behavior for  $d\sigma^{(1+2)}$

$$\begin{aligned} \frac{d\sigma^{(1+2)}}{dzd^2\vec{q}_\perp}(q_\perp \sim mz_{\text{cut}}) &\simeq \frac{d\sigma^{(1)}}{dzd^2\vec{q}_\perp}(q_\perp) \\ \frac{d\sigma^{(1+2)}}{dzd^2\vec{q}_\perp}(q_\perp \gg mz_{\text{cut}}) &\simeq \frac{d\sigma^{(2)}}{dzd^2\vec{q}_\perp}(q_\perp) \end{aligned} \quad (3.19)$$

Both these regimes have a common resummation factor (resumming double and single logarithms of  $Q/mz_{\text{cut}}$ ) which we can be ignored since it does not affect the shape of the distribution. Turning off of the resummation in region 2 is achieved using profile scales in  $\mu_{cs}$  as shown in figure 3. At the same time we probe for higher order corrections using scale variations about the central profile by a factor of two and one half.

## 4 Comparison with simulation

In this section we compare our LL+LO result against monte-carlo simulations. For simplicity we consider the process  $e^+e^- \rightarrow b + X$ , where we groom with soft-drop the hemisphere which the  $b$ -quark is found. The two hemispheres are identified by the plane perpendicular to the thrust axis [96]. We measure the energy fraction,  $z$ , and the transverse momentum,  $q_\perp$ , of the heavy-quark with respect to the groomed-jet axis.<sup>5</sup>

For the simulations we used Madgraph [97] for generating the LO hard process  $e^+e^- \rightarrow b + \bar{b}$  and then the partonic shower is implemented in PYTHIA 8 [98, 99]. For the analysis, FASTJET [100] is used along with the corresponding add-ons from FASTJET-contrib for imposing soft-drop grooming.

The kinematic variables we choose for the comparison are for center of mass energy  $\sqrt{s} = 100 \text{ GeV}$  and for the grooming parameter  $z_{\text{cut}} = 0.1$ . In figure 4 we compare the LL+LO results against the simulation for four different values of the energy fraction:  $z = 0.75, 0.80, 0.85, 0.89$ . We find good agreement with the simulation for most of the range of  $q_\perp$  and within the theoretical uncertainty. The uncertainty for the analytic result is obtain by varying the jet and soft scales,  $\mu_B$  and  $\mu_{cs}$  by a factor of 1/2 and 2 around their canonical values. We note that for small transverse momenta,  $q_\perp < 2 \text{ GeV}$ , the scale variation results in a large uncertainty. This is a result of the small scale  $mz_{\text{cut}} \sim 0.5 \text{ GeV}$  that dominates the uncertainty band in that region.

Although according to our result, the shape of the transverse momentum distribution does not explicitly depend on the hard scale  $Q$ , at larger values of  $q_\perp \sim Q(1-z)$  power corrections of the form  $q_\perp/Q(1-z)$  become important. These terms are not captured by the EFTs used here. To test these observation we compare our analytic result against the simulation for  $\sqrt{s} = 100$  and  $200 \text{ GeV}$ . The results are shown in figure 5 for the cases  $z = 0.75$  and  $z = 0.85$ . We find that indeed the partonic transverse momentum spectrum is independent of the value of  $Q$  and in agreement with our predictions away from the region  $q_\perp \sim Q(1-z)$ .

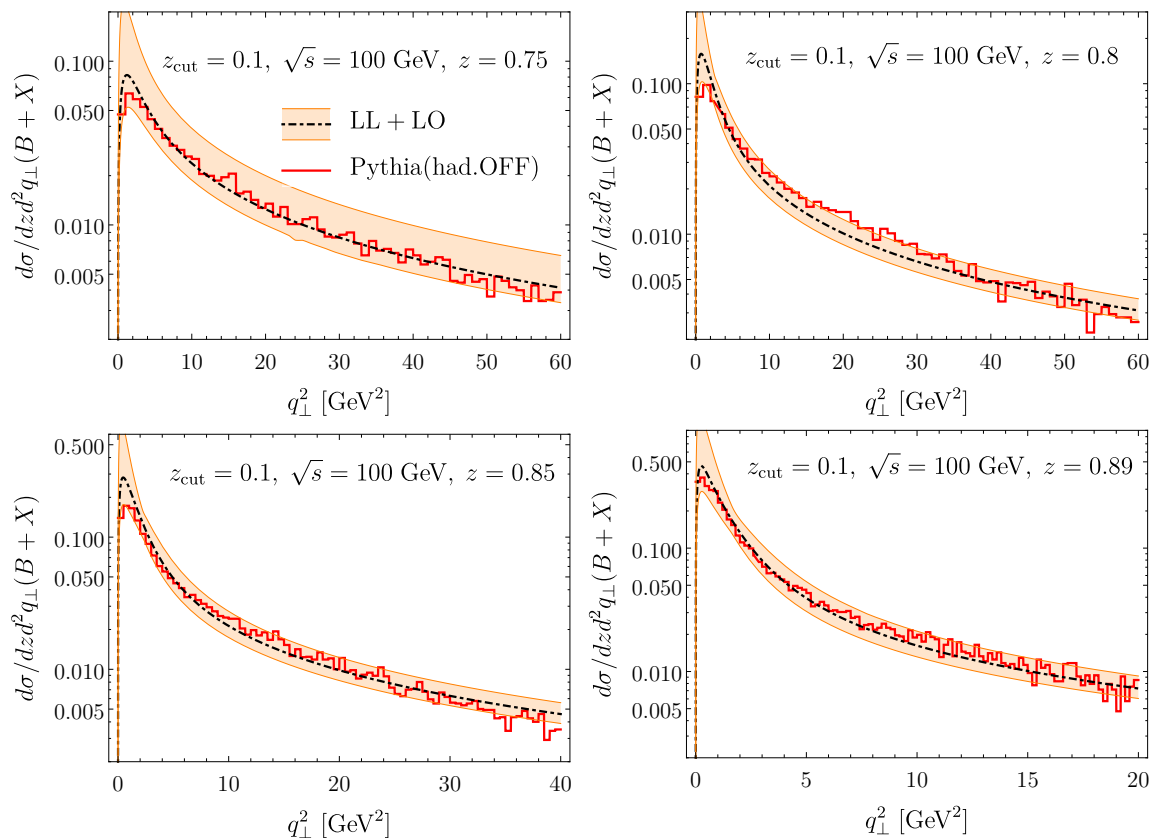
The simulation results eventually vanish at  $q_\perp^{\text{max}} = E_J(1-z)$ . In this region the heavy meson recoils against soft radiation close to the hemisphere boundary. To describe the cross section in this region one can construct an EFT with the following hierarchy:

$$Q \gg Q(1-z) \sim q_\perp \gg m. \tag{4.1}$$

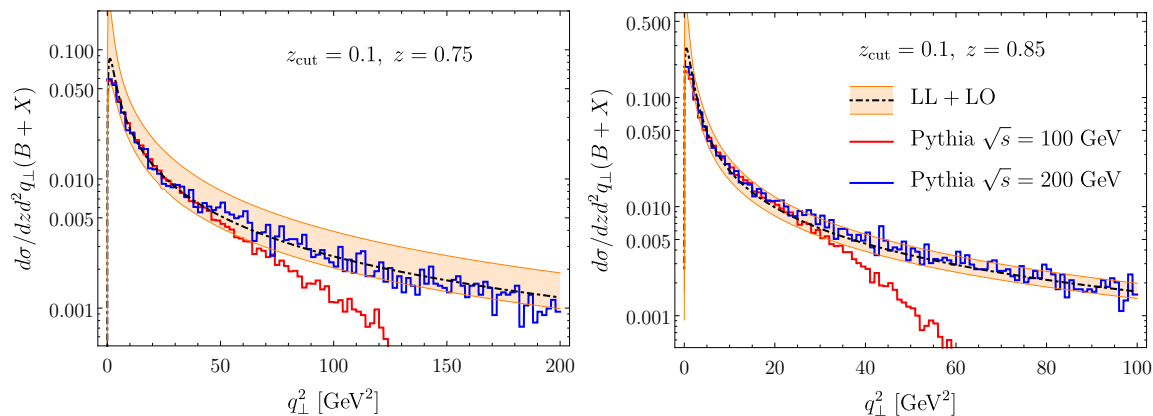
In this region the cross section is mostly insensitive to the grooming procedure and there are two relevant modes that contribute to the measurements. The global soft mode,  $p_s^\mu \sim Q(1-z)(1, 1, 1) \sim q_\perp(1, 1, 1)$  which now contribute both to the energy and transverse momentum and the ultra-collinear mode,  $p_{uc} \sim m(1-z)(Q/m, m/Q, 1)$  which only contributes to the energy measurement. Note that the naive fixed order EFT result diverge in that region. The correct behavior of the cross section can be reproduced within the EFT only at NLL' (or higher logarithmic accuracy) as described in ref. [101]. However, we do not investigate this region further in this paper since we do not go beyond a one loop calculation. At

---

<sup>5</sup>In this case the groomed-jet axis is defined as the direction of the total three-momenta of all particles in the corresponding hemisphere that pass the grooming procedure.



**Figure 4.** Comparison of the analytical result (with error bands) at LL+LO accuracy with parton level PYTHIA with a jet energy  $E_J = Q/2 \sim 50 \text{ GeV}$ .



**Figure 5.** Comparison with parton level PYTHIA for different values of jet energy  $E_J = Q/2$ . As the jet energy increases, the EFT is valid till a higher value of  $q_\perp$ . This plot vividly demonstrates the  $Q$  independence of the shape at low  $q_\perp$ .

the same time, this far tail region will be very hard to measure and hence will not be of significant phenomenological importance.

## 5 Non-perturbative corrections

In order to probe the non-perturbative corrections, we lower the virtuality of the modes from  $q_\perp$  to  $\Lambda_{\text{QCD}}$ . Since the HQET ultra-collinear mode has the lowest virtuality in our scale hierarchy, this function will contribute to the dominant non-perturbative effects. The HQET jet operator in region 1 is defined as eq. (A.4)

$$B_+^{(\perp)} = \langle 0 | \bar{h}_{v_+} W_n \delta(q^+ - (1 - \hat{\Theta}_{SD}) \mathcal{P}_X^-) \delta^2(q_\perp - (1 - \hat{\Theta}_{SD}) \mathcal{P}_\perp) | B + X \rangle \langle B + X | W_n^\dagger H_{v_+} | 0 \rangle \quad (5.1)$$

Taking the Laplace and Fourier transform with respect to  $q^+$  and  $\vec{q}_\perp$  we can write

$$B_+^{(\perp)} = \langle 0 | \bar{h}_{v_+} W_n e^{-s(1 - \hat{\Theta}_{SD}) p_X^-} e^{\vec{b} \cdot \vec{q}_\perp (1 - \hat{\Theta}_{SD})} | B + X \rangle \langle B + X | W_n^\dagger h_{v_+} | 0 \rangle \quad (5.2)$$

We consider the case when  $p_X^-, q_\perp \sim \Lambda_{\text{QCD}}$ , which then induces power corrections of the form  $s\Lambda_{\text{QCD}}, b\Lambda_{\text{QCD}}$ . Expanding out and keeping the leading order term, we have

$$B_+^{(0)} = \langle 0 | \bar{h}_{v_+} W_n | B + X \rangle \langle B + X | W_n^\dagger h_{v_+} | 0 \rangle \quad (5.3)$$

which is simply a normalization factor. Lets look at the first non trivial power correction. We have two contributions, one of which is from the expansion in  $s\Lambda_{\text{QCD}}$ .

$$P_1 = -s p_X^- \langle 0 | \bar{h}_{v_+} W_n | B + X_{SD} \rangle \langle B + X_{SD} | W_n^\dagger h_{v_+} | 0 \rangle \quad (5.4)$$

where  $X_{SD}$  indicates that this operator exists only for the radiation that passes soft-drop. To establish some type of scaling universality, we do a Lorentz transformation to the rest frame of the heavy quark. The Wilson line remains unchanged while the velocity of the heavy quark now becomes  $v \sim (1, 0, 0, 0)$  This simply gives us

$$P_1 = -s Q/m \langle 0 | \bar{h}_v W_n \mathcal{P}_X^- | B + X_{SD} \rangle \langle B + X_{SD} | W_n^\dagger h_v | 0 \rangle \quad (5.5)$$

The matrix element  $\langle 0 | \bar{h}_v W_n \mathcal{P}^- | B + X_{SD} \rangle \langle B + X_{SD} | W_n^\dagger h_v | 0 \rangle$  is independent of the energy of the jet or the mass of the heavy quark and has dimensions of energy. It is however, dependent on the soft-drop condition and hence is sensitive to the value of  $z_{\text{cut}}$ . Typically we would expect this object to have value  $\sim \Lambda_{\text{QCD}}$ . In order to go deep into the non-perturbative region  $Q(1-z) < \Lambda_{\text{QCD}}$ , in principle, we need to have an all orders description of the non-perturbative corrections. The usual way to deal with this is to put in some type of a model shape function that captures the dominant non-perturbative physics. Any such model would also have the constraint that the leading power correction should be of the form eq. (5.5). Here we consider a simple exponential model which correctly recovers the leading power correction.

$$\tilde{f}_{\text{np}}(s) = e^{-sQ\Lambda/m} \quad (5.6)$$

Taking the inverse Laplace transform we get

$$f_{\text{np}}(q^+) = \delta\left(q^+ - Q\frac{\Lambda}{m}\right) \quad (5.7)$$

which depends on a single parameter  $\Lambda$  that needs to be fitted from simulation/experiment. The differential spectrum at the hadronic level can then be given as a convolution of this shape function with the partonic jet function.

$$B_+^{(\perp)}(E_J(1-z))\Big|_{\text{had.}} = B_+^{(\perp)}(q^+) \otimes_z f(q^+) = B_+^{(\perp)}(E_J(1-z-\Lambda/m)) \quad (5.8)$$

which implements a simple shift in  $z$ , so that the hadronized spectrum is same as the partonic one at a lower value of  $z$ . More sophisticated models exist in literature [102, 103] which smoothen out the delta function about its central value of  $y = \Lambda/m$ , but have essentially have the same effect.

Figure 6 shows the comparison of hadronized PYTHIA with this simple shift model. The blue curves are the parton level distributions, while the red ones are hadron level. We have set the value of  $\Lambda \sim 0.2$  GeV, in order to have a good match with PYTHIA at lower values of  $z$  ( $z \sim 0.75$ ). Note that the parameter  $\Lambda$  found here is smaller than the corresponding  $\lambda$  found in ref. [102]. This is due to the fact that, in our analysis  $\Lambda$  depends on soft-drop which suppresses contribution from the wide angle radiation. We see that the agreement with PYTHIA worsens in at higher  $z$  for very low  $q_\perp$ . This is mainly because of the fact that the process of grooming and hadronization do not commute. Looking the  $z$  spectrum we see that the number of events at the parton level which just fail soft-drop is very large (This is to be contrasted with a gluon initiate jet where the cross section drops to zero at such high  $z$  values). On hadronization, many of these partons acquire enough energy to pass soft-drop, thus increasing the total number of events at low  $q_\perp$  at the hadronized level. All of these events are, of course, missed by the analytical calculation, which implements hadronization corrections only on those events that pass soft-drop at the partonic level.

We can do a similar boost for the power correction from  $q_\perp$  which gives us

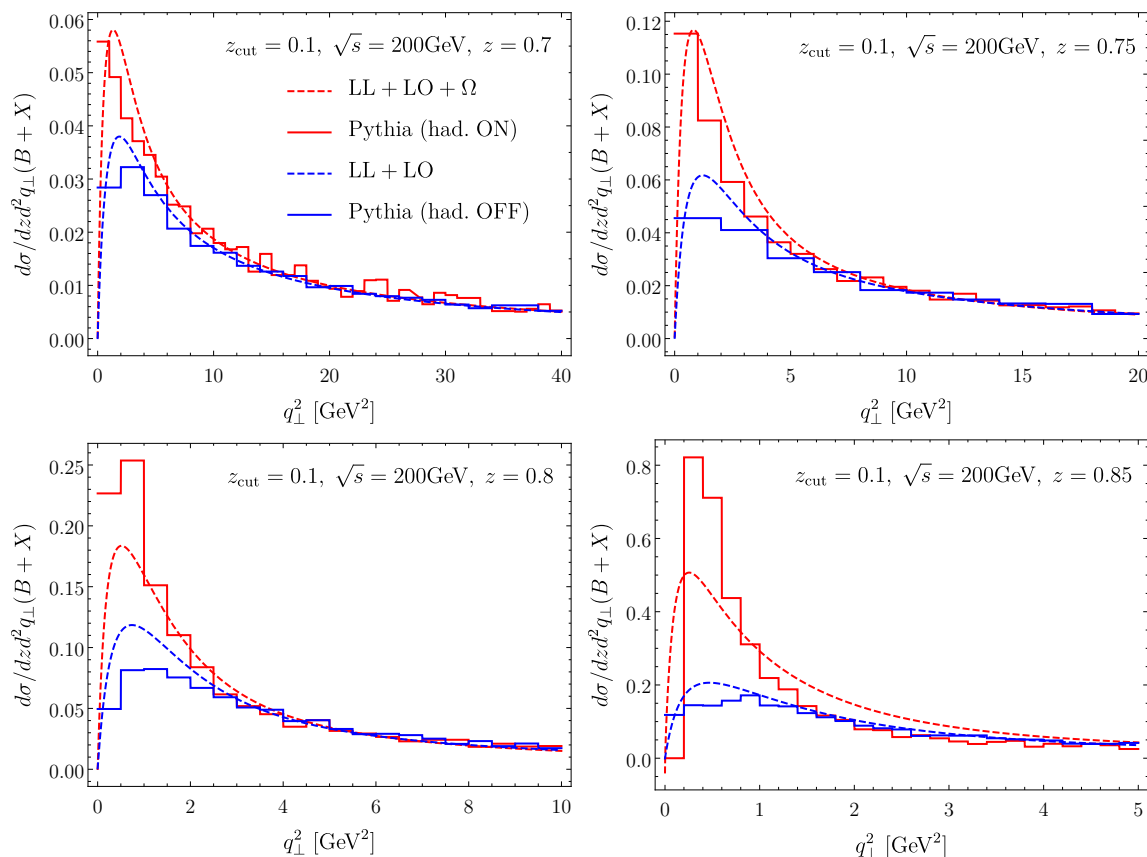
$$-b^2 \langle 0 | \bar{h}_v W_n \mathcal{P}_{X_\perp}^2 | B + X \rangle \langle B + X | W_n^\dagger h_v | 0 \rangle \quad (5.9)$$

where we have used rotational invariance. In the same vein as the power corrections in  $z$ , we can implement an exponential (in this case a Gaussian) model to implement the non-perturbative corrections. We do not explicitly implement these correction in this paper, since the perturbative error bands in the very low  $q_\perp$  region are large and we are not yet sensitive enough to the non-perturbative effects to extract out a meaningful parameter.

## 6 Conclusion

We present the transverse momentum spectrum for a heavy hadron identified in a groomed jet. We work in a regime where the heavy hadron carries most of the energy of the jet. The grooming is implemented with a soft-drop grooming procedure. The results presented are for the case of a  $b$  quark initiated jet in  $e^+e^-$  collision, although this can be easily extended to the case of  $pp$  collisions. The grooming gives us a significant advantage in that we are insensitive to corrections due to non-global logarithms or pile-up.

We work in a hierarchy  $E_J \gg E_J(1-z) \sim E_J z_{\text{cut}} \gg m \gg \Lambda_{\text{QCD}}$ , where  $E_J$  is the energy of the jet,  $z$ , the fraction of the jet energy carried by the heavy hadron, while  $z_{\text{cut}}$



**Figure 6.** Hadronization effects in PYTHIA versus the simplified prescription of eq. (5.8). The results are shown for various values of the momentum fraction, particularly we choose  $z = 0.7, 0.75, 0.8,$  and  $0.85$ . We find that, as expected, the simple shift works better for values away from the energy threshold,  $1 - z_{\text{cut}}$ .

is the grooming parameter.  $m$  is the mass of the heavy quark. We identify two distinct regimes for the transverse momentum  $q_{\perp} \sim m(1-z) \ll m$  and  $E_J(1-z) \gg q_{\perp} \gg m(1-z)$ .

We formulate two separate factorization theorems in these two regimes to capture the dominant contribution to the cross section at leading power in this hierarchy. We resum the large logarithms in  $q_{\perp}$  at Leading Log accuracy and smoothly match the cross section in the two regimes of transverse momentum. For the low  $q_{\perp}$  regime, the shape of the distribution is given entirely by the fixed order cross section and there are no large logarithms (in  $q_{\perp}$  or  $(1-z)$ ) that require resummation.

Comparing with partonic PYTHIA, we obtain an excellent agreement in the regime where the EFT is valid. There is deviation in the far tail region  $q_{\perp} \sim E_J(1-z)$ , where the cross section becomes sensitive to the boundary of the jet. From our analytic calculation, we observe that the shape of the  $q_{\perp}$  spectrum is independent of the energy of the jet over most of the  $q_{\perp}$  range, again, deviation from this behavior is in the far tail where the angle  $q_{\perp}/(E_J(1-z))$  of the radiation recoiling against the heavy quark nears the edge of the jet.

We obtain analytical expressions for the scaling behavior of the leading non-perturbative power corrections to the cross section. These leading power corrections are incorporated in a model implementing a simple shift in  $z$ . Comparing the  $q_\perp$  region, we again obtain good agreement with PYTHIA except for the very low  $q_\perp$  region. As explained in section 5, this discrepancy, especially at high values of  $z$  is due to the large number of events that pass soft-drop after hadronization, which are missing from the analytical calculation.

While our calculation only captures the leading-logarithmic (LL) behavior of the cross section, uncertainty reduction can be obtained by a higher order computation. This will require a two loop computation of the HQET jet functions in each regime of  $q_\perp$ . Especially in the low  $q_\perp$  region, the error band is large since the fixed order cross section evaluated at the scale  $mz_{\text{cut}} \sim 0.5 \text{ GeV}$  is essentially dictating the shape of the transverse momentum distribution.

We propose that this observable can be used as a probe of heavy quark TMD fragmentation function. At the same time, we can also use this measurement as a jet substructure observable since it is sensitive to the radiation pattern in the jet.

## Acknowledgments

We would like to thank Duff Neill and Emanuele Mereghetti for several useful discussions during the course of this project. This work was supported by the U.S. Department of Energy through the Office of Science, Office of Nuclear Physics under Contract DE-AC52-06NA25396 and by an Early Career Research Award, through the LANL/LDRD Program, and within the framework of the TMD Topical Collaboration.

## A One loop results

### A.1 GSoft

The global soft function is defined as the following matrix element of Wilson lines

$$S_G(z_{\text{cut}}) = \frac{1}{N_C} \text{Tr} \langle 0 | T \{ Y_n Y_{\bar{n}} \} \hat{\Theta}_{SD} \Theta_R \bar{T} \{ Y_n Y_{\bar{n}} \} | 0 \rangle \quad (\text{A.1})$$

$\hat{\Theta}_{SD}$  denotes the soft drop groomer.  $\Theta_R$  imposes the hemisphere jet constraint. We require that the global soft modes fail soft drop. From literature [44], we can write down the result for the one loop singular piece

$$S_G = 1 + \frac{\alpha_s C_F}{\pi} \ln^2 \left( \frac{\mu}{Q z_{\text{cut}}} \right) \quad (\text{A.2})$$

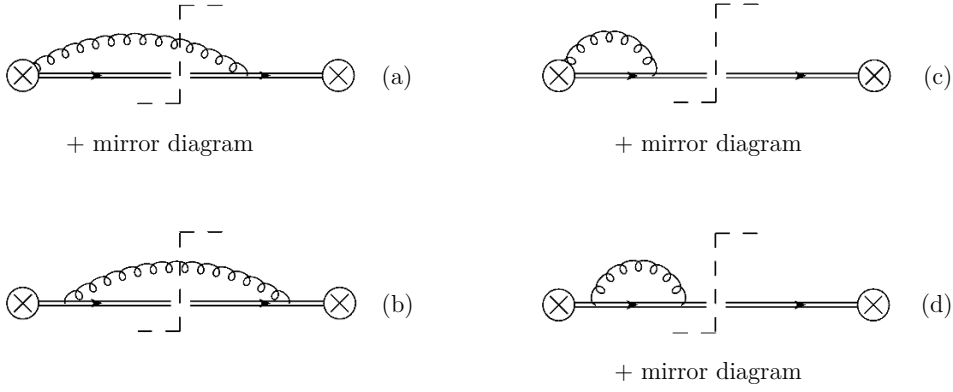
So the natural scale for this function is  $\mu_{gs} = Q z_{\text{cut}}$ .

### A.2 BHQET

#### A.2.1 Region 1: $q_\perp \sim m(1 - z)$

The residual momentum that make up the modes of this region (known as ultra collinear modes) scales as

$$k^\mu \sim \Gamma(m/Q, Q/m, 1) \quad (\text{A.3})$$



**Figure 7.** One loop diagrams for the HQET jet function.

where  $\Gamma = m(1 - z)$  is the IR scale for this EFT ( $m$  being the hard scale). In this regime  $q_\perp \sim m(1 - z)$  so that the jet function contributes to both measurements. Therefore the boosted HQET jet function is defined as

$$B_+^{(\perp)} = \frac{1}{\mathcal{N}} \langle 0 | \bar{h}_{v_+} W_n \mathcal{M}^{SD}(q^+, \vec{q}_\perp) W_n^\dagger h_{v_+} | 0 \rangle. \quad (\text{A.4})$$

where  $v_+ = (m/Q, Q/m, 0_\perp)$  is the velocity of the boosted heavy quark and  $\mathcal{M}^{SD}$  the measurement operator

$$\mathcal{M}^{SD}(q^+, \vec{q}_\perp) = \delta(q^+ - (1 - \hat{\Theta}_{SD})\mathcal{P}^-) \delta^2(\vec{q}_\perp - (1 - \hat{\Theta}_{SD})\vec{\mathcal{P}}_\perp) \quad (\text{A.5})$$

At one loop there are two real emission diagrams (diagrams (a) and (b) in figure 7).

$$\begin{aligned} B_{+,a}^{(\perp)} &= 2g^2 C_F \tilde{\mu}^{2\epsilon} \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2) \delta(\bar{n} \cdot k - q^+) \delta^2(\vec{q}_\perp - \vec{k}_\perp)}{v_+ \cdot k \bar{n} \cdot k} \theta(\bar{n} \cdot k - Qz_{\text{cut}}) \\ &\quad + 2g^2 C_F \tilde{\mu}^{2\epsilon} \delta(q^+) \delta^2(\vec{q}_\perp) \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2)}{v_+ \cdot k \bar{n} \cdot k} \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\ &\equiv B_{+,a-1}^{(\perp)} + B_{+,a-2}^{(\perp)} \end{aligned} \quad (\text{A.6})$$

$$B_{+,a-1}^{(\perp)} = \frac{2g^2 C_F}{(2\pi)^3} \frac{1}{q^+} \frac{1}{(q^+)^2 m^2/Q^2 + q_\perp^2} \theta(q^+ - Qz_{\text{cut}}) \quad (\text{A.7})$$

$$\begin{aligned} B_{+,a-2}^{(\perp)} &= 2g^2 C_F \tilde{\mu}^{2\epsilon} \delta(q^+) \delta^2(\vec{q}_\perp) \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2)}{v_+ \cdot k \bar{n} \cdot k} \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\ &= -\frac{g^2 C_F}{8\pi^2} \left( \frac{\tilde{\mu}}{mz_{\text{cut}}} \right)^{2\epsilon} \delta(q^+) \delta^2(\vec{q}_\perp) \frac{\Gamma[\epsilon]}{\epsilon} \end{aligned} \quad (\text{A.8})$$

Expanding out all the pieces and keeping the singular terms, we get

$$B_{+,a}^{(\perp)} = \frac{\alpha_s C_F}{\pi^2} \frac{1}{q^+} \frac{1}{(q^+)^2 m^2/Q^2 + q_\perp^2} \theta(q^+ - Qz_{\text{cut}}) - \frac{\alpha_s C_F}{\pi} \ln^2 \left( \frac{\mu}{mz_{\text{cut}}} \right) \quad (\text{A.9})$$

which reproduces the correct cusp anomalous dimension. We have one more diagram

$$\begin{aligned}
 B_{+,b}^{(\perp)} &= -g^2 C_F \int d^d k \frac{\delta(k^2)(v_+)^2 \delta(q^+ - \bar{n} \cdot k) \delta^2(\vec{q}_\perp - \vec{k}_\perp)}{(v_+ \cdot k)^2} \theta(\bar{n} \cdot k - Qz_{\text{cut}}) \\
 &\quad - g^2 C_F \delta(q^+) \delta^2(\vec{q}_\perp) \int d^d k \frac{\delta(k^2)(v_+)^2}{(v_+ \cdot k)^2} \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\
 &= -\frac{2g^2 C_F m^2}{(2\pi)^3 Q^2} q^+ \left( \frac{1}{m^2/Q^2(q^+)^2 + q_\perp^2} \right)^2 \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\
 &\quad + \frac{g^2 C_F}{8\pi^2} \left( \frac{\tilde{\mu}}{mz_{\text{cut}}} \right)^{2\epsilon} \frac{1}{\epsilon}
 \end{aligned} \tag{A.10}$$

The two corresponding virtual contributions from diagrams (c) and (d) in figure 7, are scaleless and disappear in dimensional regularization. If we combine all the finite terms we have,

$$B_+^{(\perp)} = \delta(q^+) \delta^{(2)}(\vec{q}_\perp) + \frac{\alpha_s C_F}{\pi} \left\{ \frac{1}{\pi} \frac{q_\perp^2}{q^+} \frac{\theta(q^+ - Qz_{\text{cut}})}{((q^+)^2 m^2/Q^2 + q_\perp^2)^2} - \ln^2 \left( \frac{\mu}{mz_{\text{cut}}} \right) + \ln \left( \frac{\mu}{mz_{\text{cut}}} \right) \right\} \tag{A.11}$$

which tells us the natural scale for this function  $\mu_B = \mu z_{\text{cut}}$ . The corresponding anomalous dimension as defined in eq. (3.13) is,

$$\gamma_B^{(\perp)} = \frac{\alpha_s C_F}{\pi} \left\{ -2 \ln \left( \frac{\mu}{mz_{\text{cut}}} \right) + 1 \right\} \tag{A.12}$$

### A.2.2 Region 2: $q_\perp \gg m(1-z)$

In this case, the HQET mode does not contribute to the measurement of the transverse momentum but otherwise has an identical definition to the previous case.

$$B_+ = \frac{1}{N} \langle 0 | \bar{h}_{v_+} W_n \delta(q^+ - (1 - \hat{\Theta}_{SD}) \mathcal{P}^-) W_n^\dagger h_{v_+} | 0 \rangle. \tag{A.13}$$

The real and virtual diagrams remain the same as before. The result for diagram (a) is

$$\begin{aligned}
 B_{+,a} &= 2g^2 C_F \tilde{\mu}^{2\epsilon} \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2) \delta(\bar{n} \cdot k - q^+) \bar{n} \cdot v_+}{v_+ \cdot k \bar{n} \cdot k} \theta(\bar{n} \cdot k - Qz_{\text{cut}}) \\
 &\quad + 2g^2 C_F \tilde{\mu}^{2\epsilon} \delta(q^+) \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2) \bar{n} \cdot v_+}{v_+ \cdot k \bar{n} \cdot k} \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\
 &= 2g^2 C_F \left( \frac{\tilde{\mu}}{(m/Q)q^+} \right)^{2\epsilon} \frac{\pi}{(2\pi)^3} \frac{\theta(q^+ - Qz_{\text{cut}})}{q^+} \Gamma[\epsilon] \\
 &\quad - \frac{g^2 C_F}{8\pi^2} \delta(q^+) \left( \frac{\tilde{\mu}}{mz_{\text{cut}}} \right)^{2\epsilon} \frac{\Gamma[\epsilon]}{\epsilon}
 \end{aligned} \tag{A.14}$$

and for diagram (b)

$$\begin{aligned}
 B_{+,b} &= -g^2 C_F \int d^d k \frac{\delta(k^2)(v_+)^2 \delta(q^+ - \bar{n} \cdot k) \theta(\bar{n} \cdot k - Qz_{\text{cut}})}{(v_+ \cdot k)^2} \\
 &\quad - g^2 C_F \delta(q^+) \int d^d k \frac{\delta(k^2)(v_+)^2}{(v_+ \cdot k)^2} \theta(Qz_{\text{cut}} - \bar{n} \cdot k) \\
 &= -\frac{g^2 C_F \pi \theta(q^+ - Qz_{\text{cut}})}{(2\pi)^3 q^+} + \frac{g^2 C_F}{8\pi^2} \left( \frac{\tilde{\mu}}{mz_{\text{cut}}} \right)^{2\epsilon} \frac{1}{\epsilon}
 \end{aligned} \tag{A.15}$$

As before the virtual contributions are scale-less and thus vanish in pure dimensional regularization. The sum of all finite terms from the real diagrams is

$$B_+ = \delta(q_+) + \frac{\alpha_s C_F}{\pi} \left\{ \left[ 2 \ln \left( \frac{\mu Q}{m q^+} \right) - 1 \right] \frac{\theta(q^+ - Q z_{\text{cut}})}{q^+} + \left[ -\ln^2 \left( \frac{\mu}{m z_{\text{cut}}} \right) + \ln \left( \frac{\mu}{m z_{\text{cut}}} \right) \right] \delta(q^+) \right\} \quad (\text{A.16})$$

The corresponding anomalous dimension as defined in eq. (3.1) is

$$\gamma_B = \frac{\alpha_s C_F}{\pi} \left\{ -2 \ln \left( \frac{\mu}{m z_{\text{cut}}} \right) + 1 + 2 \frac{\theta(q^+ - Q z_{\text{cut}})}{q^+} \right\} \quad (\text{A.17})$$

which dictates the natural scale of this function to be again  $\mu_B = z_{\text{cut}}$ .

### A.3 HQETm

We can borrow the HQET matching coefficient from literature [104]. The result is known to two loops but we only require the one loop calculation in our paper.

$$H_+(m) = \frac{\alpha_s C_F}{\pi} \left( \ln^2 \left( \frac{\mu}{m} \right) + \frac{1}{2} \ln \left( \frac{\mu}{m} \right) \right) \quad (\text{A.18})$$

which gives AD

$$\gamma_+ = 2 \frac{\alpha_s C_F}{\pi} \ln \left( \frac{\mu}{m} \right) + \frac{\alpha_s C_F}{2\pi} \quad (\text{A.19})$$

This function obviously lives at the scale  $\mu_m = m$ .

### A.4 CSoft

The collinear-soft function is not affected by the mass of the heavy quark. Hence it is the same as in a massless jet. The collinear-soft function is defined by the matrix element

$$S_C(z_{\text{cut}}, q_\perp) = \frac{1}{N_C} \text{Tr} \langle 0 | T \left( U_n^\dagger W_t \right) \mathcal{M}^{SD}(q^+, \vec{q}_\perp) \bar{T} \left( W_t^\dagger U_n \right) | 0 \rangle \quad (\text{A.20})$$

The collinear-soft modes only contribute to the measurement if they pass soft-drop, which is implemented by the  $\hat{\Theta}_{SD}$  term. Then the one-loop collinear-soft function is,

$$\begin{aligned} S_{C,a} &= 4g^2 C_F \tilde{\mu}^{2\epsilon} \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2) \delta(\vec{n} \cdot k - q^+) \delta^2(\vec{q}_\perp - \vec{k}_\perp)}{n \cdot k \bar{n} \cdot k} \theta(\vec{n} \cdot k - Q z_{\text{cut}}) \\ &\quad + 4g^2 C_F \tilde{\mu}^{2\epsilon} \delta(q^+) \delta^2(\vec{q}_\perp) \int \frac{d^d k}{(2\pi)^{d-1}} \frac{\delta^+(k^2)}{n \cdot k \bar{n} \cdot k} \theta(Q z_{\text{cut}} - \vec{n} \cdot k) \\ &\equiv S_{C,a-1} + S_{C,a-2} \end{aligned} \quad (\text{A.21})$$

$S_{C,a-2}$  is scaleless and disappears in dimensional regularization.

$$S_{C,a-1} = 2g^2 C_F \frac{1}{q_\perp^2} \left( \frac{\tilde{\mu}}{q_\perp} \right)^{2\epsilon} \frac{1}{(2\pi)^3} \frac{\theta(q^+ - Q z_{\text{cut}})}{q^+} \quad (\text{A.22})$$

Expanding in  $\epsilon$  and keeping only the finite contributions we get the renormalized collinear-soft function in  $\overline{\text{MS}}$

$$S_C = \delta(q^+) \delta^{(2)}(\vec{q}_\perp) + \frac{2\alpha_s C_F}{\pi} \mathcal{L}_0(q_\perp, \mu) \frac{\theta(q^+ - Qz_{\text{cut}})}{q^+} \quad (\text{A.23})$$

where  $\mathcal{L}_0$  is a distribution function defined as

$$\mathcal{L}_0(q_\perp, \mu) = \frac{1}{2\pi} \frac{1}{\mu^2} \left[ \frac{\mu^2}{q_\perp^2} \right]_+ \quad (\text{A.24})$$

The properties of these functions are collected in [3]. For our purposes, we need the relation

$$\mu \frac{d}{d\mu} \mathcal{L}_0(q_\perp, \mu) = -\delta^2(\vec{q}_\perp) \quad (\text{A.25})$$

and the corresponding anomalous dimension defined by eq. (3.1) (in impact parameter space) is

$$\gamma_{cs} = -2 \frac{\alpha_s C_F}{\pi} \frac{\theta(q^+ - Qz_{\text{cut}})}{q^+} \quad (\text{A.26})$$

## B Outline of factorization

In this appendix we briefly discuss the factorization formulae in the presence of soft-drop grooming. We first discuss the generic result of eq. (2.3) and later we focus on the subsequent refactorizations of the jet function in eqs. (2.4) and (2.5) for regions 1 and 2 respectively. Here we only give the outline of the relevant proof for factorization. For the reader interested in a more rigorous approach we suggest reading first section 3 of ref. [105] and section 4 of ref. [106]. We start from the full theory cross section,

$$\frac{d\sigma}{dz d^2\vec{q}_\perp} = \frac{1}{\mathcal{N}} \int dx d^2 p_{h\perp} \Big|_{p_h^2 = m^2} \sum_X \left| \langle 0 | j^\mu \mathcal{M}^{SD}(q^+, \vec{q}_\perp) | X h(x, \vec{p}_{h\perp}) \rangle L_\mu \right|^2 \delta^{(4)}(Q^\mu - p_{Xh}^\mu). \quad (\text{B.1})$$

$j^\mu$  is the electromagnetic current for quarks,  $\mathcal{M}^{SD}$  implements the Soft-drop groomed measurement.  $L^\mu$  is the wavefunction term for  $e^+e^-$  initial state. Since we are interested in the generic kinematic region  $Q \gg Q(1-z) \sim Qz_{\text{cut}} \gg |\vec{q}_\perp|$  only soft and collinear modes will contribute to the final state radiation. We thus match to SCET which is the appropriate EFT in this region of phase space. The matching process consists of two major steps:

- Match the full theory currents  $j^\mu$  onto the SCET current operators,

$$j^\mu(x) = \sum_i C_i^\mu \cdot \bar{\chi}_{\bar{n}} \Gamma_i \chi_n(x) \quad (\text{B.2})$$

where  $C_i$  are the Wilson short distance matching coefficients.

- factorize the Hilbert final states onto the subspace of relevant kinematics,

$$|X + h(z, \vec{p}_{h\perp})\rangle \rightarrow |X_{\text{reco.}}\rangle |X_{\text{coll.}} + h(z, \vec{p}_{h\perp})\rangle |X_{\text{soft}}\rangle \quad (\text{B.3})$$

where  $X_{\text{recol}}$  is the radiation recoiling to the hemisphere jet of interest while  $h$  is the final state meson that is identified in the jet. Also since we are considering only the threshold region the heavy meson is also collinear and hence inhabits the same Hilbert space as  $X_{\text{coll}}$ .

It is important to note that soft radiation has transverse momentum parametrically larger than the heavy meson and thus, from the kinetic constraints, doomed to fail the soft-drop grooming. Also important, the recoil radiation, which is collinear along the opposite hemisphere axis, is decoupled from collinear modes in the SCET Lagrangian. In addition, the soft modes can be decoupled from collinear radiation through the BPS field redefinition,

$$\begin{aligned}\chi_n(x) &= Y_n^\dagger(x)\chi_n^{(0)}(x), \\ \bar{\chi}_n(x) &= \bar{\chi}_n^{(0)}(x)Y_n(x),\end{aligned}\tag{B.4}$$

where  $\chi_n^{(0)}$  is the uncoupled quark field. Pushing all of the normalization constants, leptonic tensors, Wilson matching coefficients, and recoiling radiation matrix elements into one factor,  $\sigma_0(E_J)$  we can write the cross section as follows:

$$\begin{aligned}\frac{d\sigma}{dzd^2\vec{q}_\perp} &= \sigma_0(E_J) \times \frac{1}{N_C} \text{Tr}\langle 0|T\{Y_n Y_{\bar{n}}\}\hat{\Theta}_{SD}\Theta_R\bar{T}\{Y_n Y_{\bar{n}}\}|0\rangle \\ &\times \int dx d^2p_{h\perp} \Big|_{p_h^2=m^2} \sum_{X_{\text{coll}}} \left| \langle 0|\chi_n \mathcal{M}^{SD}(q^+, \vec{q}_\perp)|X_{\text{coll}}h(x, \vec{p}_{h\perp})\rangle \right|^2 \delta^{(4)}(p_{jet}^\mu - p_{X_{\text{coll}},h}^\mu).\end{aligned}\tag{B.5}$$

defining then the second term in the first line as the global soft function and the second line as the jet function we retrieve eq. (2.3).

For further refactorization of the jet function we need to match the collinear fields of SCET onto SCET<sub>M</sub> and subsequently onto bHQET. This process is discussed in detail in ref. [107] and is effectively achieved through the following replacement,

$$\chi_n \rightarrow C_+ W_n^\dagger h_{v_+}\tag{B.6}$$

Here  $C_+$  is the amplitude level matching coefficient from SCET<sub>M</sub> to bHQET.  $W_n^\dagger$  is the Wilson line obtained by decoupling the ultra-collinear mode (which describes the IR fluctuations around a boosted heavy quark) of bHQET from all the other sectors. Defining  $H_+ = C_+ C_+^*$  we have the refactorization of the jet function to the hard function  $H_+$  and the bHQET jet function as described in eq. (2.4)

If a hierarchy exists between the scales  $q_\perp$  and  $m(1-z)$  then we are in region 2 and a second collinear mode (referred to as the collinear-soft mode) is relevant to measurement which needs to be included in the effective theory. The collinear-soft mode has the following scaling,

$$p_{sc}^\mu \sim (Q(1-z), q_\perp^2/Q(1-z), q_\perp)\tag{B.7}$$

The matching from SCET onto the new effective theory may be achieved with the following replacement,

$$\chi_n \rightarrow C_+ W_t^\dagger W_n^\dagger h_{v_+}\tag{B.8}$$

where the collinear Wilson line  $W_t$  consist of collinear-soft fields and the matching coefficients  $C_+$  are the same as in matching onto bHQET. At this stage the heavy quark field and the collinear-soft modes are still coupled at the level of the Lagrangian. One needs to perform a BPS-like field redefinition to achieve further refactorization. This is similar to what is done in SCET<sub>+</sub> in ref. [106], i.e.,

$$W_n^\dagger h_{v_+} = U_n W_n^{(0)\dagger} h_{v_+}^{(0)} \tag{B.9}$$

where  $W_n^{(0)\dagger}$  and  $h_{v_+}^{(0)}$  are decoupled from the collinear-soft fields. To proceed we need to apply power-counting to the measurement operator  $\mathcal{M}^{SD}$  and more particularly to the operators  $\mathcal{P}^-$  and  $\mathcal{P}_\perp$ . While the contribution of ultra-collinear and collinear-soft modes to the energy fraction is similar the contribution of the ultra-collinear in perpendicular component is suppressed,

$$\begin{aligned} \mathcal{P}^- &\sim \mathcal{P}_{cs}^- + \mathcal{P}_{uc}^- \\ \mathcal{P}_\perp &\sim \mathcal{P}_{\perp cs} \end{aligned} \tag{B.10}$$

With these modifications is then trivial to show the refactorization of the bHQET jet function into an ultra-collinear component and the collinear soft function as shown in eq. (2.5).

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

## References

- [1] J. Collins, *Foundations of perturbative QCD*, Cambridge University Press, Cambridge, U.K., (2013).
- [2] M.G. Echevarria, A. Idilbi and I. Scimemi, *Factorization theorem for Drell-Yan at low  $q_T$  and transverse momentum distributions on-the-light-cone*, *JHEP* **07** (2012) 002 [[arXiv:1111.4996](https://arxiv.org/abs/1111.4996)] [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [3] J.-Y. Chiu, A. Jain, D. Neill and I.Z. Rothstein, *A formalism for the systematic treatment of rapidity logarithms in quantum field theory*, *JHEP* **05** (2012) 084 [[arXiv:1202.0814](https://arxiv.org/abs/1202.0814)] [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [4] A.V. Manohar and I.W. Stewart, *The zero-bin and mode factorization in quantum field theory*, *Phys. Rev.* **D 76** (2007) 074002 [[hep-ph/0605001](https://arxiv.org/abs/hep-ph/0605001)] [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [5] H. Georgi and H.D. Politzer, *Quark decay functions and heavy hadron production in QCD*, *Nucl. Phys.* **B 136** (1978) 445 [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [6] R.K. Ellis, H. Georgi, M. Machacek, H.D. Politzer and G.G. Ross, *Perturbation theory and the parton model in QCD*, *Nucl. Phys.* **B 152** (1979) 285 [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [7] J.C. Collins and D.E. Soper, *Parton distribution and decay functions*, *Nucl. Phys.* **B 194** (1982) 445 [[INSPIRE](https://inspirehep.net/literature/1000000)].
- [8] J.C. Collins, D.E. Soper and G.F. Sterman, *Factorization of hard processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [[hep-ph/0409313](https://arxiv.org/abs/hep-ph/0409313)] [[INSPIRE](https://inspirehep.net/literature/1000000)].

- [9] J.C. Collins and D.E. Soper, *Back-to-back jets in QCD*, *Nucl. Phys. B* **193** (1981) 381 [Erratum *ibid.* **B 213** (1983) 545] [INSPIRE].
- [10] R. Bain, L. Dai, A. Hornig, A.K. Leibovich, Y. Makris and T. Mehen, *Analytic and Monte Carlo studies of jets with heavy mesons and quarkonia*, *JHEP* **06** (2016) 121 [arXiv:1603.06981] [INSPIRE].
- [11] M. Procura and I.W. Stewart, *Quark fragmentation within an identified jet*, *Phys. Rev. D* **81** (2010) 074009 [Erratum *ibid.* **D 83** (2011) 039902] [arXiv:0911.4980] [INSPIRE].
- [12] A. Jain, M. Procura and W.J. Waalewijn, *Parton fragmentation within an identified jet at NNLL*, *JHEP* **05** (2011) 035 [arXiv:1101.4953] [INSPIRE].
- [13] A. Jain, M. Procura, B. Shotwell and W.J. Waalewijn, *Fragmentation with a cut on thrust: predictions for B-factories*, *Phys. Rev. D* **87** (2013) 074013 [arXiv:1207.4788] [INSPIRE].
- [14] C.W. Bauer and E. Mereghetti, *Heavy quark fragmenting jet functions*, *JHEP* **04** (2014) 051 [arXiv:1312.5605] [INSPIRE].
- [15] M. Dasgupta, F. Dreyer, G.P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039 [arXiv:1411.5182] [INSPIRE].
- [16] M. Baumgart, A.K. Leibovich, T. Mehen and I.Z. Rothstein, *Probing quarkonium production mechanisms with jet substructure*, *JHEP* **11** (2014) 003 [arXiv:1406.2295] [INSPIRE].
- [17] T. Kaufmann, A. Mukherjee and W. Vogelsang, *Hadron fragmentation inside jets in hadronic collisions*, *Phys. Rev. D* **92** (2015) 054015 [arXiv:1506.01415] [INSPIRE].
- [18] Z.-B. Kang, F. Ringer and I. Vitev, *Jet substructure using semi-inclusive jet functions in SCET*, *JHEP* **11** (2016) 155 [arXiv:1606.07063] [INSPIRE].
- [19] Z.-B. Kang, F. Ringer and I. Vitev, *The semi-inclusive jet function in SCET and small radius resummation for inclusive jet production*, *JHEP* **10** (2016) 125 [arXiv:1606.06732] [INSPIRE].
- [20] Y.-T. Chien, Z.-B. Kang, F. Ringer, I. Vitev and H. Xing, *Jet fragmentation functions in proton-proton collisions using soft-collinear effective theory*, *JHEP* **05** (2016) 125 [arXiv:1512.06851] [INSPIRE].
- [21] R. Bain, Y. Makris and T. Mehen, *Transverse momentum dependent fragmenting jet functions with applications to quarkonium production*, *JHEP* **11** (2016) 144 [arXiv:1610.06508] [INSPIRE].
- [22] L. Dai, C. Kim and A.K. Leibovich, *Fragmentation of a jet with small radius*, *Phys. Rev. D* **94** (2016) 114023 [arXiv:1606.07411] [INSPIRE].
- [23] L. Dai, C. Kim and A.K. Leibovich, *Fragmentation to a jet in the large  $z$  limit*, *Phys. Rev. D* **95** (2017) 074003 [arXiv:1701.05660] [INSPIRE].
- [24] Z.-B. Kang, X. Liu, F. Ringer and H. Xing, *The transverse momentum distribution of hadrons within jets*, *JHEP* **11** (2017) 068 [arXiv:1705.08443] [INSPIRE].
- [25] Z.-B. Kang, A. Prokudin, F. Ringer and F. Yuan, *Collins azimuthal asymmetries of hadron production inside jets*, *Phys. Lett. B* **774** (2017) 635 [arXiv:1707.00913] [INSPIRE].
- [26] Z.-B. Kang, F. Ringer and W.J. Waalewijn, *The energy distribution of subjets and the jet shape*, *JHEP* **07** (2017) 064 [arXiv:1705.05375] [INSPIRE].

- [27] J.R. Forshaw, M.H. Seymour and A. Siodmok, *On the breaking of collinear factorization in QCD*, *JHEP* **11** (2012) 066 [[arXiv:1206.6363](#)] [[INSPIRE](#)].
- [28] S. Catani, D. de Florian and G. Rodrigo, *Space-like (versus time-like) collinear limits in QCD: is factorization violated?*, *JHEP* **07** (2012) 026 [[arXiv:1112.4405](#)] [[INSPIRE](#)].
- [29] T.C. Rogers and P.J. Mulders, *No generalized TMD-factorization in hadro-production of high transverse momentum hadrons*, *Phys. Rev. D* **81** (2010) 094006 [[arXiv:1001.2977](#)] [[INSPIRE](#)].
- [30] M. Zeng, *Drell-Yan process with jet vetoes: breaking of generalized factorization*, *JHEP* **10** (2015) 189 [[arXiv:1507.01652](#)] [[INSPIRE](#)].
- [31] J.R. Gaunt, *Glauber gluons and multiple parton interactions*, *JHEP* **07** (2014) 110 [[arXiv:1405.2080](#)] [[INSPIRE](#)].
- [32] I.Z. Rothstein and I.W. Stewart, *An effective field theory for forward scattering and factorization violation*, *JHEP* **08** (2016) 025 [[arXiv:1601.04695](#)] [[INSPIRE](#)].
- [33] A. Banfi, G. Marchesini and G. Smye, *Away from jet energy flow*, *JHEP* **08** (2002) 006 [[hep-ph/0206076](#)] [[INSPIRE](#)].
- [34] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett. B* **512** (2001) 323 [[hep-ph/0104277](#)] [[INSPIRE](#)].
- [35] A.J. Larkoski, I. Moult and B. Nachman, *Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning*, [arXiv:1709.04464](#) [[INSPIRE](#)].
- [36] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft drop*, *JHEP* **05** (2014) 146 [[arXiv:1402.2657](#)] [[INSPIRE](#)].
- [37] M. Dasgupta, A. Fregoso, S. Marzani and A. Powling, *Jet substructure with analytical methods*, *Eur. Phys. J. C* **73** (2013) 2623 [[arXiv:1307.0013](#)] [[INSPIRE](#)].
- [38] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [[arXiv:1307.0007](#)] [[INSPIRE](#)].
- [39] A.J. Larkoski, I. Moult and D. Neill, *Analytic boosted boson discrimination at the Large Hadron Collider*, [arXiv:1708.06760](#) [[INSPIRE](#)].
- [40] A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, *Extracting a short distance top mass with light grooming*, [arXiv:1708.02586](#) [[INSPIRE](#)].
- [41] S. Marzani, L. Schunk and G. Soyez, *A study of jet mass distributions with grooming*, *JHEP* **07** (2017) 132 [[arXiv:1704.02210](#)] [[INSPIRE](#)].
- [42] M. Dasgupta, A. Powling, L. Schunk and G. Soyez, *Improved jet substructure methods: Y-splitter and variants with grooming*, *JHEP* **12** (2016) 079 [[arXiv:1609.07149](#)] [[INSPIRE](#)].
- [43] C. Frye, A.J. Larkoski, M.D. Schwartz and K. Yan, *Precision physics with pile-up insensitive observables*, [arXiv:1603.06375](#) [[INSPIRE](#)].
- [44] C. Frye, A.J. Larkoski, M.D. Schwartz and K. Yan, *Factorization for groomed jet substructure beyond the next-to-leading logarithm*, *JHEP* **07** (2016) 064 [[arXiv:1603.09338](#)] [[INSPIRE](#)].
- [45] A.J. Larkoski, S. Marzani and J. Thaler, *Sudakov safety in perturbative QCD*, *Phys. Rev. D* **91** (2015) 111501 [[arXiv:1502.01719](#)] [[INSPIRE](#)].

- [46] M. Dasgupta, L. Schunk and G. Soyez, *Jet shapes for boosted jet two-prong decays from first-principles*, *JHEP* **04** (2016) 166 [[arXiv:1512.00516](#)] [[INSPIRE](#)].
- [47] Y.-T. Chien and I. Vitev, *Jet shape resummation using soft-collinear effective theory*, *JHEP* **12** (2014) 061 [[arXiv:1405.4293](#)] [[INSPIRE](#)].
- [48] CMS collaboration, *Search for a Higgs boson in the decay channel  $H \rightarrow ZZ^{(*)} \rightarrow q\bar{q}\ell^-\ell^+$  in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, *JHEP* **04** (2012) 036 [[arXiv:1202.1416](#)] [[INSPIRE](#)].
- [49] CMS collaboration, *Search for a Standard Model-like Higgs boson decaying into  $WW \rightarrow \ell\nu q\bar{q}$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*, *CMS-PAS-HIG-13-008*, CERN, Geneva, Switzerland, (2013).
- [50] ATLAS collaboration, *Measurement of jet charge in dijet events from  $\sqrt{s} = 8$  TeV  $pp$  collisions with the ATLAS detector*, *Phys. Rev. D* **93** (2016) 052003 [[arXiv:1509.05190](#)] [[INSPIRE](#)].
- [51] CMS collaboration, *Search for BSM  $t\bar{t}$  production in the boosted all-hadronic final state*, *CMS-PAS-EXO-11-006*, CERN, Geneva, Switzerland, (2011).
- [52] ATLAS and CMS collaborations, *Boosted top quark techniques and searches for  $t\bar{t}$  resonances at the LHC*, *J. Phys. Conf. Ser.* **452** (2013) 012034 [[INSPIRE](#)].
- [53] ATLAS and CMS collaborations, J. Pilot, *Boosted top quarks, top pair resonances and top partner searches at the LHC*, *EPJ Web Conf.* **60** (2013) 09003 [[INSPIRE](#)].
- [54] ATLAS collaboration, *Performance of boosted top quark identification in 2012 ATLAS data*, *ATLAS-CONF-2013-084*, CERN, Geneva, Switzerland, (2013).
- [55] J.D. Bjorken, *Energy loss of energetic partons in quark-gluon plasma: possible extinction of high  $p_T$  jets in hadron-hadron collisions*, FERMILAB-PUB-82-059-THY, Fermilab, Batavia, IL, U.S.A., (1982) [FERMILAB-PUB-82-059-T] [[INSPIRE](#)].
- [56] M. Gyulassy and M. Plumer, *Jet quenching in dense matter*, *Phys. Lett. B* **243** (1990) 432 [[INSPIRE](#)].
- [57] X.-N. Wang and M. Gyulassy, *Gluon shadowing and jet quenching in  $A+A$  collisions at  $\sqrt{s} = 200$  GeV*, *Phys. Rev. Lett.* **68** (1992) 1480 [[INSPIRE](#)].
- [58] ATLAS collaboration, *Measurements of the nuclear modification factor for jets in  $Pb+Pb$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with the ATLAS detector*, *Phys. Rev. Lett.* **114** (2015) 072302 [[arXiv:1411.2357](#)] [[INSPIRE](#)].
- [59] CMS collaboration, *Studies of jet quenching using isolated-photon+jet correlations in  $Pb+Pb$  and  $pp$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*, *Phys. Lett. B* **718** (2013) 773 [[arXiv:1205.0206](#)] [[INSPIRE](#)].
- [60] CMS collaboration, *Observation and studies of jet quenching in  $Pb+Pb$  collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV*, *Phys. Rev. C* **84** (2011) 024906 [[arXiv:1102.1957](#)] [[INSPIRE](#)].
- [61] ATLAS collaboration, *Observation of a centrality-dependent dijet asymmetry in lead-lead collisions at  $\sqrt{s_{NN}} = 2.77$  TeV with the ATLAS detector at the LHC*, *Phys. Rev. Lett.* **105** (2010) 252303 [[arXiv:1011.6182](#)] [[INSPIRE](#)].
- [62] CMS collaboration, *Measurement of jet fragmentation into charged particles in  $pp$  and  $Pb+Pb$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*, *JHEP* **10** (2012) 087 [[arXiv:1205.5872](#)] [[INSPIRE](#)].

- [63] CMS collaboration, *Modification of jet shapes in Pb+Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*, *Phys. Lett. B* **730** (2014) 243 [[arXiv:1310.0878](#)] [[INSPIRE](#)].
- [64] ATLAS collaboration, *Measurement of inclusive jet charged-particle fragmentation functions in Pb+Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with the ATLAS detector*, *Phys. Lett. B* **739** (2014) 320 [[arXiv:1406.2979](#)] [[INSPIRE](#)].
- [65] I. Feige, M.D. Schwartz, I.W. Stewart and J. Thaler, *Precision jet substructure from boosted event shapes*, *Phys. Rev. Lett.* **109** (2012) 092001 [[arXiv:1204.3898](#)] [[INSPIRE](#)].
- [66] M. Field, G. Gur-Ari, D.A. Kosower, L. Mannelli and G. Perez, *Three-prong distribution of massive narrow QCD jets*, *Phys. Rev. D* **87** (2013) 094013 [[arXiv:1212.2106](#)] [[INSPIRE](#)].
- [67] A.J. Larkoski, J. Thaler and W.J. Waalewijn, *Gaining (mutual) information about quark/gluon discrimination*, *JHEP* **11** (2014) 129 [[arXiv:1408.3122](#)] [[INSPIRE](#)].
- [68] M. Dasgupta, A. Powling and A. Siodmok, *On jet substructure methods for signal jets*, *JHEP* **08** (2015) 079 [[arXiv:1503.01088](#)] [[INSPIRE](#)].
- [69] J.R. Walsh and S. Zuberi, *Factorization constraints on jet substructure*, [arXiv:1110.5333](#) [[INSPIRE](#)].
- [70] A.J. Larkoski, I. Moult and D. Neill, *Power counting to better jet observables*, *JHEP* **12** (2014) 009 [[arXiv:1409.6298](#)] [[INSPIRE](#)].
- [71] A.J. Larkoski, I. Moult and D. Neill, *Building a better boosted top tagger*, *Phys. Rev. D* **91** (2015) 034035 [[arXiv:1411.0665](#)] [[INSPIRE](#)].
- [72] J. Cogan, M. Kagan, E. Strauss and A. Schwartzman, *Jet-images: computer vision inspired techniques for jet tagging*, *JHEP* **02** (2015) 118 [[arXiv:1407.5675](#)] [[INSPIRE](#)].
- [73] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images — deep learning edition*, *JHEP* **07** (2016) 069 [[arXiv:1511.05190](#)] [[INSPIRE](#)].
- [74] L.G. Almeida, M. Backović, M. Cliche, S.J. Lee and M. Perelstein, *Playing tag with ANN: boosted top identification with pattern recognition*, *JHEP* **07** (2015) 086 [[arXiv:1501.05968](#)] [[INSPIRE](#)].
- [75] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet substructure classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **93** (2016) 094034 [[arXiv:1603.09349](#)] [[INSPIRE](#)].
- [76] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban and D. Whiteson, *Jet flavor classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **94** (2016) 112002 [[arXiv:1607.08633](#)] [[INSPIRE](#)].
- [77] J.S. Conway, R. Bhaskar, R.D. Erbacher and J. Pilot, *Identification of high-momentum top quarks, Higgs bosons and W and Z bosons using boosted event shapes*, *Phys. Rev. D* **94** (2016) 094027 [[arXiv:1606.06859](#)] [[INSPIRE](#)].
- [78] J. Barnard, E.N. Dawe, M.J. Dolan and N. Rajcic, *Parton shower uncertainties in jet substructure analyses with deep neural networks*, *Phys. Rev. D* **95** (2017) 014018 [[arXiv:1609.00607](#)] [[INSPIRE](#)].
- [79] S. Fleming, A.H. Hoang, S. Mantry and I.W. Stewart, *Top jets in the peak region: factorization analysis with NLL resummation*, *Phys. Rev. D* **77** (2008) 114003 [[arXiv:0711.2079](#)] [[INSPIRE](#)].

- [80] L. Dai, C. Kim and A.K. Leibovich, *Heavy quark jet fragmentation*, *JHEP* **09** (2018) 109 [[arXiv:1805.06014](#)] [[INSPIRE](#)].
- [81] C.W. Bauer, S. Fleming and M.E. Luke, *Summing Sudakov logarithms in  $B \rightarrow X_s \gamma$  in effective field theory*, *Phys. Rev. D* **63** (2000) 014006 [[hep-ph/0005275](#)] [[INSPIRE](#)].
- [82] C.W. Bauer, S. Fleming, D. Pirjol and I.W. Stewart, *An effective field theory for collinear and soft gluons: heavy to light decays*, *Phys. Rev. D* **63** (2001) 114020 [[hep-ph/0011336](#)] [[INSPIRE](#)].
- [83] C.W. Bauer and I.W. Stewart, *Invariant operators in collinear effective theory*, *Phys. Lett. B* **516** (2001) 134 [[hep-ph/0107001](#)] [[INSPIRE](#)].
- [84] C.W. Bauer, D. Pirjol and I.W. Stewart, *Soft collinear factorization in effective field theory*, *Phys. Rev. D* **65** (2002) 054022 [[hep-ph/0109045](#)] [[INSPIRE](#)].
- [85] C.W. Bauer, S. Fleming, D. Pirjol, I.Z. Rothstein and I.W. Stewart, *Hard scattering factorization from effective field theory*, *Phys. Rev. D* **66** (2002) 014017 [[hep-ph/0202088](#)] [[INSPIRE](#)].
- [86] A.V. Manohar and M.B. Wise, *Heavy quark physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **10** (2000) 1 [[INSPIRE](#)].
- [87] S.D. Ellis and D.E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160 [[hep-ph/9305266](#)] [[INSPIRE](#)].
- [88] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions*, *Nucl. Phys. B* **406** (1993) 187 [[INSPIRE](#)].
- [89] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [[hep-ph/9707323](#)] [[INSPIRE](#)].
- [90] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998–1999*, pg. 270 [[hep-ph/9907280](#)] [[INSPIRE](#)].
- [91] M. Wobisch, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at  $\sqrt{s} = 300$  GeV*, Ph.D. thesis, Tech. Hochsch., Aachen, Germany, (2000) [[INSPIRE](#)].
- [92] G. Salam,  $E_t^\infty$  scheme, unpublished.
- [93] D. Bertolini, T. Chan and J. Thaler, *Jet observables without jet algorithms*, *JHEP* **04** (2014) 013 [[arXiv:1310.7584](#)] [[INSPIRE](#)].
- [94] A.J. Larkoski, D. Neill and J. Thaler, *Jet shapes with the broadening axis*, *JHEP* **04** (2014) 017 [[arXiv:1401.2158](#)] [[INSPIRE](#)].
- [95] Y. Makris, D. Neill and V. Vaidya, *Probing transverse-momentum dependent evolution with groomed jets*, *JHEP* **07** (2018) 167 [[arXiv:1712.07653](#)] [[INSPIRE](#)].
- [96] E. Farhi, *A QCD test for jets*, *Phys. Rev. Lett.* **39** (1977) 1587 [[INSPIRE](#)].
- [97] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [[arXiv:1405.0301](#)] [[INSPIRE](#)].
- [98] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [[hep-ph/0603175](#)] [[INSPIRE](#)].

- [99] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [[arXiv:0710.3820](#)] [[INSPIRE](#)].
- [100] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [[arXiv:1111.6097](#)] [[INSPIRE](#)].
- [101] A.J. Larkoski, I. Moult and D. Neill, *Toward multi-differential cross sections: measuring two angularities on a single jet*, *JHEP* **09** (2014) 046 [[arXiv:1401.4458](#)] [[INSPIRE](#)].
- [102] M. Fickinger, S. Fleming, C. Kim and E. Mereghetti, *Effective field theory approach to heavy quark fragmentation*, *JHEP* **11** (2016) 095 [[arXiv:1606.07737](#)] [[INSPIRE](#)].
- [103] Z. Ligeti, I.W. Stewart and F.J. Tackmann, *Treating the b quark distribution function with reliable uncertainties*, *Phys. Rev. D* **78** (2008) 114014 [[arXiv:0807.1926](#)] [[INSPIRE](#)].
- [104] A.H. Hoang, A. Pathak, P. Pietrulewicz and I.W. Stewart, *Hard matching for boosted tops at two loops*, *JHEP* **12** (2015) 059 [[arXiv:1508.04137](#)] [[INSPIRE](#)].
- [105] S.D. Ellis, C.K. Vermilion, J.R. Walsh, A. Hornig and C. Lee, *Jet shapes and jet algorithms in SCET*, *JHEP* **11** (2010) 101 [[arXiv:1001.0014](#)] [[INSPIRE](#)].
- [106] C.W. Bauer, F.J. Tackmann, J.R. Walsh and S. Zuberi, *Factorization and resummation for dijet invariant mass spectra*, *Phys. Rev. D* **85** (2012) 074006 [[arXiv:1106.6047](#)] [[INSPIRE](#)].
- [107] S. Fleming, A.H. Hoang, S. Mantry and I.W. Stewart, *Jets from massive unstable particles: top-mass determination*, *Phys. Rev. D* **77** (2008) 074010 [[hep-ph/0703207](#)] [[INSPIRE](#)].