Journal of the Korean Physical Society, Vol. 74, No. 5, March 2019, pp. 439∼458

# Factorized Groomed Jet Mass Distribution in Inclusive Jet Processes

Junegone CHAY*
*Department of Physics, Korea University, Seoul 02841, Korea*

Chul KIM†
*Institute of Convergence Fundamental Studies and School of Liberal Arts, Seoul National University of Science and Technology, Seoul 01811, Korea and Pittsburgh Particle Physics Astrophysics and Cosmology Center (PITT PACC), Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*

(Received 6 September 2018)

We consider the factorized groomed jet mass distribution in inclusive jet processes by using a modified mass drop tagger (mMDT), corresponding to a soft drop with angular exponent $\beta = 0$. A grooming procedure is implemented rather than tagging in the sense that grooming always returns a groomed jet, while tagging dose not return a jet when a single particle remains after tagging. We find that the grooming procedure makes the jet mass distribution infrared safe and that only ultraviolet divergences appear in each factorized part. The groomed jet mass distributions are investigated over a wide range of the jet mass considering various limits on the jet mass variable $\rho = M_J^2/(p_T^J R)^2$ and the grooming cut $y_c$. Appropriate effective theories in different kinematic regions are employed to resum large logarithms, in which the analysis in the region $\rho \sim y_c \ll 1$ is included due to the different type of factorization. The analytic computation of the factorized groomed jet mass distribution is presented by resumming the large logarithms in the jet mass, and $y_c$. Numerically, the effect of the resummation is notably enhanced, compared with the calculation at next-to-leading order, and nonglobal logarithms are estimated to be small.

PACS numbers: 12.38.−t, 12.38.Cy, 13.87.−a
Keywords: Jet, Grooming, Mass drop tagger, Soft drop
DOI: 10.3938/jkps.74.439

## I. INTRODUCTION

Study of jet substructures has become more important in the era of LHC. At high energy, heavy particles such as the $W$, $Z$, Higgs bosons or the top quark are boosted, and the decay products become energetic and collinear, producing jets. The signal jets from heavy particles will have different jet substructures compared to jets produced by quantum chromodynamics (QCD). In order to obtain clearer information on the signal jets, we must filter out the background. The background consists of multiple parton scattering, pileup, *etc.*, called underlying events, which will affect the invariant jet mass significantly. Once we understand how the background can be controlled, we can give a more precise description of the signal jet. On the other hand, also of importance is how the background can be reduced in the jets produced by QCD, which also become the background for the signal jets. Here, we consider how to tame the background and probe the substructure of the QCD jets, which will be the starting point for studying the substructure of the signal jets.

For the analysis of jet substructure, many methods that remove wide-angle soft radiation have been proposed. For example, the mass drop tagger (MDT) [1] and its modified version (mMDT) [2] were developed, and recently a soft drop [3] has been suggested, which includes the mMDT as a special case with angular exponent $\beta = 0$. The main goal is to take away soft particles, which are likely to come from background, and to obtain information on the hard jets, which can be related to the partons participating in the hard scattering.

One of the most important observables in jet substructure is the jet mass distribution [4–11], and the plain jet mass distribution with small jet radius using the soft-collinear effective theory (SCET) [12–15] was considered in Refs. 16–20. Studies of the groomed jet mass distribution have been rapidly increasing with recent experimental analyses (See, for example, Ref. 21). In Refs. 2,3,22, the analytical calculation for the jet mass distributions with grooming was extensively investigated in QCD. Recently, the groomed jet mass distribution was studied through the resummation of large logarithms [23–27].

*E-mail: chay@korea.ac.kr
†E-mail: chul@seoultech.ac.kr

 

The analyses were performed to next-to-leading logarithmic (NLL) accuracy, except for Ref. 23,24, in which the groomed jet mass distribution was obtained to next-to-next-to-leading logarithmic (NNLL) accuracy by using a soft drop (mMDT as a special case) in SCET.

We consider the groomed jet mass distribution in inclusive jet cross sections, in which an energetic parton can be fragmented to the observed jet. The jet radius $R$ is assumed to be small, and the jet is described basically by collinear interactions.[1] In this process, we investigate the groomed jet mass distribution from the fragmenting jet functions (FJFs) [29–31]. Here, the FJFs can provide detailed information on the jet substructures. The advantage of considering the groomed jet mass distribution in the FJFs is that the dependences on the jet mass and the grooming parameter reside only in this part in the scattering cross section. Because the FJFs describe the properties of the final-state particles, they can be studied independent of the scattering processes. That is, we can probe the jet mass distribution in hadron-hadron scattering, as well as in $e^+e^-$ annihilation, with a slight modification of the kinematic variables.

In the construction of the jet mass distribution, one encounters the issue of "grooming" versus "tagging". Though no general agreement on the exact definition on grooming and tagging exists, grooming implies that it always returns an output jet, while that is not the case with tagging [2]. Theoretically, this also becomes an issue even at lowest order in which a jet consists of two partons. In tagging, if the energy-cut criterion is satisfied, the jet is tagged with the original jet mass as the tagged mass. Otherwise, no jet contributes to the tagged jet mass distribution. The removal of the jet when the tagging criterion is not satisfied renders the jet mass distribution infrared (IR) sensitive as the jet mass $M_J$ approaches zero. In grooming, when the grooming criterion is not satisfied, the final jet consisting of a single energetic parton is included in the jet mass distribution. This contribution is concentrated where the jet mass $M_J$ is zero. Therefore, adopting grooming means that we add the contribution of the $\delta(M_J^2)$ part to the groomed jet mass distribution when the grooming criterion is not satisfied.

In our paper, we choose the groomer in computing the jet mass distributions. In fact, we show that the groomer is theoretically better because, after factorization, each factorized part of the groomed jet mass distribution turns out to be IR safe. This IR safety enables us to resum large logarithms via the renormalization group (RG) equation. We emphasize that the RG equation is meaningful only when IR safety exists, or at least, IR and ultraviolet (UV) divergences are separated so that nonperturbative IR dynamics does not mix with UV behavior, as in the case of the parton distribution functions. Moreover, the normalization of the groomed jet mass distributions can be smoothly connected to the plain jet mass distribution without grooming. Furthermore, after the normalization, the groomed jet mass distribution becomes independent of the renormalization scale.

Introducing the dimensionless variable $\rho = M_J^2/(p_T^J R)^2$, we probe the groomed jet mass distribution over a wide range of $\rho$ compared to the grooming parameter $y_c$ in the mMDT. Here, $p_T^J$ is the jet transverse momentum (before grooming) relative to the beam axis. Over the entire range of the jet masses, distinct kinematic regions that have their own characteristics exist: (i) $\rho \sim y_c \sim \mathcal{O}(1)$. If we take the limit $y_c \to 0$, the region corresponds to $y_c \ll \rho \sim \mathcal{O}(1)$, which is called the tail region. (ii) $\rho \ll y_c \sim \mathcal{O}(1)$. (iii) $\rho \sim y_c \ll 1$, which we call the midrange region. Finally (iv) $\rho \ll y_c \ll 1$, which we call the peak region. Note that $y_c$ will be fixed at 0.1 in the numerical analysis, but it can be regarded as small or on the of order 1 compared to the value of $\rho$.

The tail region corresponds to the ungroomed case, and the resummation of large logarithms of $\rho$ and $y_c$ ($z_{\rm cut}$) near the peak region has been considered in Refs. 23, 24, 27. Here, we newly include the resummation in the midrange region with $\rho \sim y_c \ll 1$, in which a different factorization structure is obtained for the resummation of $\ln\rho$ and $\ln y_c$. In all these regions, various collinear modes with different scaling behavior exist, resulting in different types of factorization. We employ appropriate effective theories to compute the factorized parts and resum the large logarithms of $\ln\rho$ or $\ln y_c$. Also, we confirm that all the factorized functions are IR safe and contain only an UV divergence when the grooming procedure is applied. Based on the results, we are able to describe the behavior of the groomed jet mass distribution over a wide range of $\rho$, say, from $\mathcal{O}(10^{-5})$ to $\mathcal{O}(1)$, with different forms of factorization.

The organization of the paper is as follows: In Sec. II, the plain jet mass distribution from the FJFs before grooming is briefly described, and the basis for defining the groomed jet mass distribution is given. In Sec. III, we explain how the grooming procedure is applied to the FJFs and define the groomed jet mass distribution. The theoretical implementation of grooming is described in detail, especially how to obtain IR safe results. In Sec. IV, we identify important modes in different regions, employ the relevant effective theories, and compute the factorized parts of the groomed jet mass distribution to next-to-leading order (NLO) in $\alpha_s$. In Sec. V, we resum the large logarithms that appear in the midrange and the peak regions to NLL accuracy. In Sec. VI, a numerical analysis is performed for the groomed jet mass distributions over the entire range of the jet mass by interpolating the resummed results in various regions. The effect of the nonglobal logarithms is also estimated. In

[1] Strictly speaking, our theoretical approach in this paper holds only for $R \ll 1$, and the legitimate choice of $R$ for a phenomenological study would be $R = 0.1 - 0.4$. However, the small $R$ approximation is known to work well even in the case with $R \gtrsim 0.6$ [8,28], so our analysis here can, in practice, be applied to the jet with a sizable value of $R$.

Sec. VII, we present conclusions.

In Appendix A, we list all the functions appearing in the text. In Appendix B, the relation between the $\Lambda$-distribution and the standard plus distribution functions is given. In Appendix C, details in obtaining $\tilde{\mathcal{C}}_k(y_c, Q^2, \mu)$ in case (i) and $\mathcal{S}_k^{II}(y_c^2 Q^2, \mu)$ in case (iv) are presented with the structure of the phase spaces. In Appendix D, we show the results in the midrange and the peak regions, employing a soft drop with $\beta \geq 0$.

## II. JET MASS DISTRIBUTION FROM THE JET FRAGMENTATION FUNCTION

We consider the inclusive jet cross section in $N_1 N_2 \to JX$, where $N_1$ and $N_2$ are incoming hadrons, $J$ is a jet with a small radius $R$, and $X$ denotes all the remaining particles. The scattering cross section can be written as [32]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y\mathrm{d}p_T^J} = \sum_i \int_{x_J = p_T^J/Q_T}^1 \frac{\mathrm{d}x}{x} \frac{\mathrm{d}\sigma_i(x_J/x, \mu)}{\mathrm{d}y\mathrm{d}p_T^i} D_{J/i}(x; p_T^i R, \mu), \tag{1}$$

where $\sigma_i$ is the partonic cross section in which the final-state jet is produced by parton $i$ with the maximum transverse momentum $Q_T$ with respect to the beam direction for a given rapidity $y$. The function $D_{J/i}$, which we call the fragmentation function to a jet (FFJ), describes the probability of producing the outgoing jet $J$ from the mother parton $i$ with the momentum fraction $x$. The jet cross section in Eq. (1) necessarily involves large logarithms of small $R$, which can be resummed through the DGLAP evolutions of the FFJs [32–36].

Throughout this paper, we use $\mathrm{k_T}$-type algorithms in defining the jet, which include the $\mathrm{k_T}$ [37,38], the anti-$\mathrm{k_T}$ [39], and the Cambridge/Aachen (C/A) algorithms [40]. Up to NLO in $\alpha_s$ (two particles in a jet at most), these three algorithms have the same constraint for merging into a jet with small $R$:

$$\theta < R' \begin{cases} R' = R & \text{for } e^+e^- \text{ annihilation,} \\ R' = R/\cosh y & \text{for hadron collision.} \end{cases} \tag{2}$$

Then, the typical jet scale $E_J R'$ can be written as $p_T^J R$ for hadron-hadron collisions and $E_J R$ for $e^+e^-$ annihilation.

We go further to consider the fragmentation inside the jet [34, 36, 41, 42]. To NLO in $\alpha_s$, the factorized cross section reads [36]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y\mathrm{d}p_T^J\mathrm{d}z} = \sum_{i,k} \int_{x_J}^1 \frac{\mathrm{d}x}{x} \frac{\mathrm{d}\sigma_i(x_J/x, \mu)}{\mathrm{d}y\mathrm{d}p_T^i} \times D_{J_k/i}(x; p_T^i R, \mu) D_{l/J_k}(z; p_T^J R). \tag{3}$$

Here, $l$ is the hadron or the subjet inside the jet $J_k$ initiated by the parton $k$, and $z = p_T^l/p_T^J$ is the momentum fraction of $l$ with respect to $J_k$. We call $D_{l/J_k}(z)$ the jet fragmentation function (JFF) for $l$, which is the probability for the fragmentation process $k \to l$ inside the jet $J_k$. When we consider the jet substructure, focusing on the JFF has an advantage because it is process and renormalization scale independent. The remaining product of the partonic cross section with the FFJ in Eq. (3) is scale invariant as well [36].

The FFJ and the JFF are the probabilities satisfying the momentum sum rules

$$\begin{aligned} &\sum_k \int_0^1 \mathrm{d}x x D_{J_k/i}(x; p_T^i R, \mu) = 1, \\ &\sum_l \int_0^1 \mathrm{d}z z D_{l/J_k}(z; p_T^J R) = 1. \end{aligned} \tag{4}$$

The fragmenting processes from the quark and the gluon jets in SCET can be described by

$$\begin{aligned} \tilde{D}_{l/J_q}(z, \mu) &= \sum_{X \in J} \frac{1}{2N_c z} \int d^{D-2}\mathbf{p}_\perp \\ &\times \mathrm{Tr}\langle 0|\delta\Big(\frac{p_+}{z} - \mathcal{P}_+\Big)\delta^{(D-2)}(\boldsymbol{\mathcal{P}}_\perp)\frac{\overline{n}\!\!\!/}{2}\chi_n|l(p_+, \mathbf{p}_\perp)X\rangle \\ &\times \langle l(p_+, \mathbf{p}_\perp)X|\overline{\chi}_n|0\rangle, \\ \tilde{D}_{l/J_g}(z, \mu) &= \sum_{X \in J} \frac{1}{z p_J^+ (D-2)(N_c^2-1)} \int d^{D-2}\mathbf{p}_\perp \\ &\times \mathrm{Tr}\langle 0|\delta\Big(\frac{p_+}{z} - \mathcal{P}_+\Big)\delta^{(D-2)}(\boldsymbol{\mathcal{P}}_\perp)\mathcal{B}_n^{\perp\mu,a}|l(p_+, \mathbf{p}_\perp)X\rangle \\ &\times \langle l(p_+, \mathbf{p}_\perp)X|\mathcal{B}_{n,\mu}^{\perp,a}|0\rangle, \end{aligned} \tag{5}$$

where the spacetime dimension is $D = 4 - 2\epsilon$, and $\chi_n = W_n^\dagger \xi_n$ is the collinear quark field in the $n$-lightcone direction, with the collinear Wilson line $W_n$. Also, $\mathcal{B}_n^{\perp\mu,a} = i\overline{n}^\rho g_\perp^{\mu\nu} G^b_{n,\rho\nu} \mathcal{W}_n^{ba} = i\overline{n}^\rho g_\perp^{\mu\nu} \mathcal{W}_n^{\dagger,ba} G^b_{n,\rho\nu}$ is the collinear gluon field strength tensor, where $\mathcal{W}_n$ is the Wilson line in the adjoint representation. In our convention, $p_+ \equiv \overline{n}\cdot p = p_0 + \hat{\mathbf{n}}_J \cdot \mathbf{p}$ and $p_- \equiv n \cdot p = p_0 - \hat{\mathbf{n}}_J \cdot \mathbf{p}$, where $n^2 = \overline{n}^2 = 0$, $n \cdot \overline{n} = 2$, and $\hat{\mathbf{n}}_J$ is the unit vector along the jet direction.

Note that $\tilde{D}_{l/J_k}$ and the integrated jet function $\mathcal{J}_k$ within the jet are related by

$$\sum_l \int_0^1 \mathrm{d}z z \tilde{D}_{l/J_k}(z; p_T^J R, \mu) = \mathcal{J}_k(p_T^J R, \mu). \tag{6}$$

To one loop, the integrated jet functions with the $\mathrm{k_T}$-

type algorithms are given as [41,43–45]

$$\mathcal{J}_q(p_T^J R,\mu) = 1 + \frac{\alpha_s C_F}{2\pi}\left[\frac{3}{2}\ln\frac{\mu^2}{p_T^{J2}R^2} + \frac{1}{2}\ln^2\frac{\mu^2}{p_T^{J2}R^2} + \frac{13}{2} - \frac{3\pi^2}{4}\right], \tag{7}$$

$$\mathcal{J}_g(p_T^J R,\mu) = 1 + \frac{\alpha_s C_A}{2\pi}\left[\frac{\beta_0}{2C_A}\ln\frac{\mu^2}{p_T^{J2}R^2} + \frac{1}{2}\ln^2\frac{\mu^2}{p_T^{J2}R^2} + \frac{67}{9} - \frac{23n_f}{18C_A} - \frac{3\pi^2}{4}\right], \tag{8}$$

where $\beta_0 = 11N_c/3 - 2n_f/3$, $C_A = N_c = 3$, and $n_f$ is the number of quark flavors. The JFF $D_{h/J_k}$ is normalized using $\tilde{D}_{h/J_k}$ divided by $\mathcal{J}_k$, and is given by

$$D_{l/J_k}(z;p_T^J R) = \frac{\tilde{D}_{h/J_k}(z;p_T^J R,\mu)}{\mathcal{J}_k(p_T^J R,\mu)}\ . \tag{9}$$

Because $\mathcal{J}_k$ is included in the FFJs, $D_{J_k/i}(x)$ in Eq. (3) [36], no double counting of $\mathcal{J}_k$ takes place.

In order to consider the jet mass distribution for the inclusive jet production, we focus on the jet mass distribution of the JFF, *i.e.*, the fragmenting jet function (FJF) [29–31]. The normalized FJF within the jet, $\mathcal{G}_{l/J_k}$, is defined as

$$D_{l/J_k}(z) = \int \mathrm{d}M_J^2\ \mathcal{G}_{l/J_k}(z,M_J^2). \tag{10}$$

When the momentum sum rule is applied, the plain jet mass distribution is given by [19]

$$\Phi_k^{\rm pl}(M_J^2;p_T^J R) = \sum_{l=q,\bar{q},g}\int_0^1 \mathrm{d}z z\ \mathcal{G}_{l/J_k}(z,M_J^2). \tag{11}$$

The plain jet mass distribution is scale invariant and normalized to one because it satisfies

$$\int \mathrm{d}M_J^2\,\Phi_k^{\rm pl}(M_J^2;p_T^J R) = \sum_{l=q,\bar{q},g}\int_0^1 \mathrm{d}z z\ D_{l/J_k}(z) = 1. \tag{12}$$

Applying the momentum sum rule for $z = p_T^l/p_T^J$ in Eq. (3), the factorization theorem for the cross section with the jet mass $M_J$ and the transverse jet momentum $p_T^J$ can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y\mathrm{d}p_T^J\mathrm{d}M_J^2} = \sum_{i,k}\int_{x_J}^1 \frac{\mathrm{d}x}{x}\frac{\mathrm{d}\sigma_i(x_J/x,\mu)}{\mathrm{d}y\mathrm{d}p_T^i} \times D_{J_k/i}(x;p_T^i R,\mu)\Phi_k^{\rm pl}(M_J^2;p_T^J R). \tag{13}$$

Equation (13) is our starting point for studying the groomed jet mass distribution for the inclusive jet process. From Eq. (11), the grooming condition in the mMDT can be accomplished by restricting the range of $z$ to $y_c/(1+y_c) < z < 1/(1+y_c)$. In Eq. (13), if $x$ is not too close to one and $M_J \sim p_T^J R$, the FFJs, $D_{J_k/i}$, and the jet mass distributions, $\Phi_k^{\rm pl}$, can be described solely in terms of the collinear modes scaling as $(p_+,p_-,\mathbf{p}_\perp) \sim p_T^J(1,R^2,R)$.

If the observables are sensitive to collinear-soft (csoft) radiation [16,46–48], $\Phi_k^{\rm pl}$ and $D_{J_k/i}$ can be refactorized by including the csoft interactions [19, 27, 49, 50]. For example, when $M_J \ll p_T^J R$, $\Phi_k^{\rm pl}(M_J^2)$ can be refactorized as [19]

$$\Phi_k^{\rm pl}(M_J^2 \ll p_T^{J2}R^2;p_T^J R) = \mathcal{C}_k(p_T^{J2}R^2,\mu) \times \int_0^{M_J^2}\mathrm{d}M^2 J_k(M^2;\mu)S_k(M_J^2-M^2;p_T^J R,\mu), \tag{14}$$

where the collinear functions $\mathcal{C}_k$ are given by the inverse of the integrated jet functions $\mathcal{J}_k^{-1}$ for $k = q, g$ and enter as the normalization. $J_k(M^2)$ are the "standard jet functions" introduced in Refs. 14,51–53, where $J_q$ is defined as

$$\sum_{X_n}\langle 0|\chi_n^\alpha|X_n\rangle\langle X_n|\bar{\chi}_n^\beta|0\rangle = \int\frac{d^4p}{(2\pi)^3}p_+\frac{\bar{n}\!\!\!/}{2}J_q(p^2,\mu)\delta^{\alpha\beta}, \tag{15}$$

and $J_g$ is similarly defined in terms of $\mathcal{B}_n^{\perp\mu,a}$. The functions $J_k$ are governed by the "ultracollinear" modes, which scale as $p_{\rm uc} = (p_{\rm uc}^+, p_{\rm uc}^-, \mathbf{p}_{\rm uc}^\perp) = p_T^J(1, M_J^2/p_T^{J2}, M_J/p_T^J)$.

Finally, the csoft functions $S_k$ describe the interaction of the csoft modes, which scale as

$$p_{\rm cs}^\mu = (p_{\rm cs}^+, p_{\rm cs}^-, \mathbf{p}_{\rm cs}^\perp) \sim \frac{M_J^2}{p_T^J R^2}(1,R^2,R) = \rho p_T^J(1,R^2,R). \tag{16}$$

The quark csoft function $S_q$ is defined as

$$S_q(M^2 = p_J^+\ell_-,\mu) = \frac{1}{p_J^+ N_c} \times \mathrm{Tr}\,\langle 0|Y_{\overline{n},\rm cs}^\dagger Y_{n,\rm cs}\delta(\ell_- + \Theta(R-\theta)i\partial_-)Y_{n,\rm cs}^\dagger Y_{\overline{n},\rm cs}|0\rangle, \tag{17}$$

where $Y_{n(\overline{n}),\rm cs}$ are the csoft Wilson lines, in which the soft gauge field in the soft Wilson lines [14] is replaced by the csoft gauge field. The gluon csoft function $S_g$ is expressed in terms of the csoft Wilson lines in the adjoint representation. Note that the collinear and the csoft modes are sensitive to the jet boundary characterized by the jet radius $R$ while the ultracollinear modes

are too narrow to recognize it. The large hierarchy in energy between the collinear and csoft modes may give rise to large nonglobal logarithms (NGLs) [54,55], which cause sizable uncertainty when estimating the small jet mass distributions.

## III. JET GROOMING AND THE GROOMED JET MASS DISTRIBUTION

We employ the mMDT [2] for grooming, which involves two parameters $y_c$ and $\mu$. It is prescribed as follows: For an initial jet $j$,

1. Decluster the jet $j$ into two subjets $j_1$ and $j_2$ with $m_{j_1} > m_{j_2}$ by undoing the last clustering process.
2. If a significant mass drop occurs, $m_{j_1} < \mu m_j$, and the splitting satisfies the criterion
$$\min(p_{Tj_1}^2, p_{Tj_2}^2)\frac{\Delta R_{j_1 j_2}^2}{m_j^2} > y_c, \tag{18}$$
then take $j$ to be the tagged jet.
3. Otherwise, redefine $j$ to be that of $j_1$ and $j_2$ with the larger transverse mass $m^2 + p_T^2$ and go back to step 1 (unless $j$ consists of a single particle, in which case the original jet is deemed untagged).

At leading order in which the jet consists of two partons, when the jet is declustered, each jet is massless, the mass drop condition is automatically satisfied, and the parameter $\mu$ is irrelevant. Then, if the criterion in Eq. (18) is satisfied, the jet is tagged. Otherwise, the jet does not contribute to the tagged jet mass distribution. On the other hand, a soft drop has a more generalized criterion. Compared to the mMDT condition in Eq. (18), it is given as

$$\frac{\min(p_{t1}, p_{t2})}{p_{t1} + p_{t2}} > z_{\rm cut}\Big(\frac{\theta_{12}}{R}\Big)^{\beta}, \tag{19}$$

where $z_{\rm cut}$ has the same role as $y_c$ in mMDT and $\beta$ is the angular exponent to control the dependence of the angle between the two partons, $\theta_{12} = \sqrt{\Delta y_{12}^2 + \Delta\phi_{12}^2}$. The limit $\beta \to 0$ corresponds to the mMDT. We will also use the soft drop in the midrange and the peak regions.

This is an original prescription, but we further implement the idea of "grooming", which always returns a groomed jet. This means that, when a single particle remains in the jet after grooming, we include the contribution of the single particle in the groomed jet mass distribution. That is, we include the contribution of the jet mass with $M_J = 0$. Even though the criterion for the mMDT or the soft drop is not satisfied, the remaining particles contribute to the jet mass. This also holds when only a single particle remains in the jet. Therefore, the "groomed jet mass distribution" starts from $\delta(M_J^2)$ at order $\alpha_s^0$. Also, at NLO, we include the virtual corrections to the jet mass distribution, which cancel the IR divergence in the real emissions. This makes the resultant groomed jet mass distribution IR safe even in the limit $M_J \to 0$. Because we do not drop any event as in the plain jet mass, the normalization is the same as in the plain jet mass distribution, *i.e.*,

$$\int {\rm d}M_J^2\ \Phi_{k=q,g}(y_c, M_J^2) = \int {\rm d}M_J^2\ \Phi_k^{\rm pl}(M_J^2) = 1, \tag{20}$$

where $\Phi_k(y_c, M_J^2)$ is the groomed jet mass distribution. Despite the same normalization, the dependence of the groomed jet mass distribution on $M_J$ is different from that of the plain jet mass distribution because the jet mass is different in the regions where the criterion for the mMDT in Eq. (18) or for the soft drop in Eq. (19) is not satisfied.

By applying grooming with mMDT to Eq. (11), as prescribed above, we can express the groomed jet mass distribution at NLO in $\alpha_s$ as

$$\Phi_k(y_c, M_J^2) = \sum_{l=q,\bar q,g}\Bigg[\int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z z\mathcal{G}_{l/J_k}(z, M_J^2) + \delta(M_J^2)\Big(\int_0^{z_{\rm min}} {\rm d}z + \int_{z_{\rm max}}^1 {\rm d}z\Big)\int_0^{M_{\rm max}^2} {\rm d}M^2 z\mathcal{G}_{l/J_k}(z, M^2)\Bigg], \tag{21}$$

where $z_{\rm min} = y_c/(1+y_c)$, $z_{\rm max} = 1/(1+y_c)$, and $M_{\rm max}^2 = z(1-z)p_T^{J2}R^2$. At NLO, if the criterion in Eq. (18) is not satisfied, only a single energetic particle is included. It is given by the second term in Eq. (21), which is proportional to $\delta(M_J^2)$, and the virtual corrections are included.

If the second term is discarded in Eq. (21), that corresponds to jet tagging, in which no single energetic parton is included in the final jet. We can show from explicit calculations at NLO, that the jet mass distribution in this case becomes IR sensitive as $M_J$ goes to zero. Let us consider the region $\rho \sim y_c \sim \mathcal{O}(1)$ as an example. In this region, only the collinear modes with $(p_c^+, p_c^-, \mathbf{p}_c^\perp) \sim p_T^J(1, R^2, R)$ contribute to the jet mass distribution. We obtain the "tagged jet mass distribution" as

$$\Phi_{k=q,g}^{\rm tag}(y_c, M_J^2) = \frac{\alpha_s C_k}{2\pi}\Bigg\{-\delta(M_J^2)\Big(\frac{1}{\epsilon_{\rm IR}} + \ln\frac{\mu^2}{M_c^2} + h_k(y_c)\Big) + f_k(w)\frac{\theta(M_J^2 - M_c^2)}{M_J^2} + g_k(y_c)\Big[\frac{\theta(M_c^2 - M_J^2)}{M_J^2}\Big]_{M_c^2}\Bigg\}, \tag{22}$$

where $C_k = C_F, C_A$ for $k = q, g$, $M_c^2 = p_T^{J2}R^2 y_c/(1+y_c)^2$ and $w = \sqrt{1 - 4M_J^2/(p_T^J R)^2}$. The functions $f_k(y)$, $g_k(w)$, and $h_k(y)$ are given in Eq. (A3), Eq. (A4), and

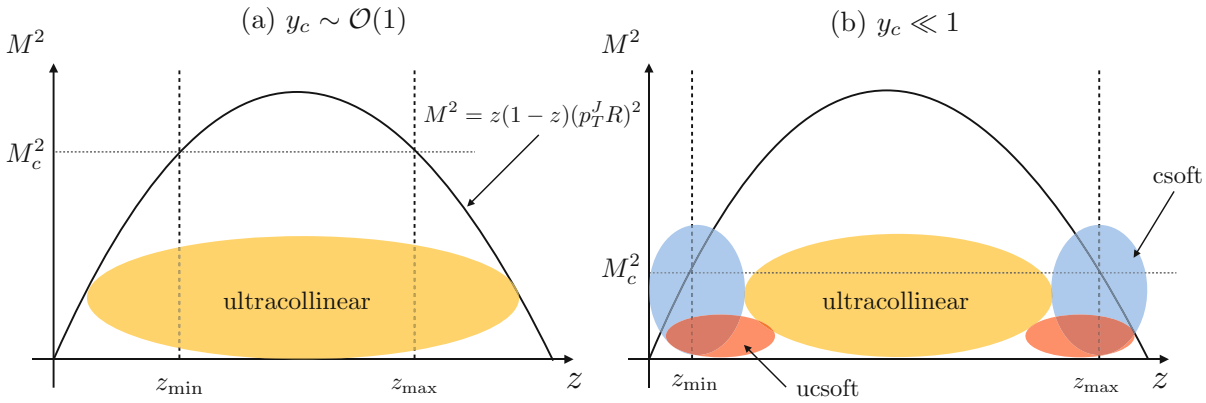


Fig. 1. (Color online) Subsets of the collinear modes are shown in $z$-$M^2$ space for (a) $y_c \sim \mathcal{O}(1)$ and (b) $y_c \ll 1$. Here, $z$ is the energy fraction of a parton, and $M^2$ is the invariant mass squared. The parabola is the jet boundary dictated by the $k_T$-type algorithm. The region between $z_{\max}$ and $z_{\min}$ under the parabola is the region satisfying the mMDT criterion, Eq. (18).

Eq. (A6), respectively. In extracting the IR divergence as $M_J^2$ goes to zero, we employ the "$\Lambda$-distribution", which is defined as

$$\int_0^{\mathcal{M}^2} \mathrm{d}M^2 [G(M^2)]_{\Lambda^2} F(M^2) = \int_0^{\mathcal{M}^2} \mathrm{d}M^2 G(M^2) F(M^2) - \int_0^{\Lambda^2} \mathrm{d}M^2 G(M^2) F(0), \quad (23)$$

for any smooth function $F(M^2)$ near $M^2 = 0$. In Eq. (22), $\Lambda^2$ has been set to $M_c^2$.

Due to the IR divergence being proportional to $\delta(M_J^2)$ in Eq. (22), the distribution cannot be normalized unless we impose a small nonzero jet mass cut, $M_{J,\mathrm{cut}}$. Then the normalization of the tagged jet mass distribution depends on the mass cut, which results in an unwanted uncertainty. For example, if $M_{J,\mathrm{cut}} \ll p_T^J R$, the normalization involves a large logarithm involving the mass cut.

On the other hand, the groomed jet mass distribution is IR safe because the second term in Eq. (21) cancels the IR divergence in the tagged distribution. The NLO result reads

$$\Phi_k(y_c, M_J^2) = \delta(M_J^2)\Big(1 - \frac{\alpha_s C_k}{2\pi} I_k(y_c)\Big) + \frac{\alpha_s C_k}{2\pi}\left\{ f_k(w) \frac{\theta(M_J^2 - M_c^2)}{M_J^2} + g_k(y_c) \Big[\frac{\theta(M_c^2 - M_J^2)}{M_J^2}\Big]_{M_c^2} \right\}, \quad (24)$$

where the functions $I_k(y)$ are given in Eq. (A1). Note that the groomed jet mass distribution in Eq. (24) is the same as that in the tagged case in Eq. (22) except for the terms proportional to $\delta(M_J^2)$. Therefore, for the jet mass distribution, the distinction between "jet grooming" and "jet tagging" is related to the issue of how to treat the jet with $M_J = 0$, *e.g.*, with a single parton in the jet.

Taking grooming over tagging has a theoretical advantage. First, we can properly normalize the jet mass distribution in Eq. (20) without imposing the mass cut. Second, because we keep the leading-order (LO) result, the nonzero $M_J$ part appears at NLO in the grooming procedure.[2] This makes factorizing the jet mass distribution convenient when we consider various small limits of $\rho$ and $y_c$. Through the NLO calculations in various limits, we can understand the factorization structure and its consistency transparently. Moreover, the resummation of the large logarithms arising from small $\rho$ and $y_c$ can be systematically performed mainly from the NLO results for each factorized part, as we will show later. In the next section, we investigate the groomed jet mass distributions in various limits of $\rho$ and $y_c$. We employ appropriate effective theories for different modes in each region to establish the factorization.

## IV. GROOMED JET MASS DISTRIBUTION IN VARIOUS REGIONS

In Eq. (24), we have considered the groomed jet mass distribution in the tail region $\rho \sim y_c \sim \mathcal{O}(1)$ without taking specific limits on $\rho = M^2/(p_T^J R)^2$ and $y_c$. In this case, we can describe the distribution only with the collinear modes scaling as $(p_c^+, p_c^-, \mathbf{p}_c^\perp) \sim p_T^J(1, R^2, R)$. In this section, we consider the cases with small jet mass for different hierarchies of $\rho$ and $y_c$. They are given as

$$(a)\ \rho \ll y_c \sim \mathcal{O}(1), \qquad (b)\ y_c \ll 1, \quad \begin{cases} \rho \sim y_c: & \text{the midrange region,} \\ \rho \ll y_c: & \text{the peak region.} \end{cases} \quad (25)$$

[2] In tagging, because the second term in Eq. (21) is removed, the LO result appears at order $\alpha_s$; hence, the NLO result is counted as $\mathcal{O}(\alpha_s^2)$. Therefore, two-loop calculations for the jet mass are needed in order to resum the large logarithms via the RG equations in the tagging procedure.

Because the jet mass is small, characteristic scales smaller than the collinear scale $\mu_c \sim p_T^J R$ exist. In fact, various subsets of the collinear modes are required to describe the dynamics at scales lower than the collinear scale.

In case (a) with $\rho \ll y_c \sim \mathcal{O}(1)$, the collinear mode cannot contribute to the small jet mass because $p_c^2 \gg M_J^2$. Instead the ultracollinear modes contribute to the jet mass, with the scaling

$$p_{\rm uc}^\mu \sim p_T^J\Big(1, \frac{M_J^2}{p_T^{J2}}, \frac{M_J}{p_T^J}\Big) = p_T^J(1, \rho R^2, \sqrt{\rho}R), \qquad (26)$$

where $p_{\rm uc}^2 \sim M_J^2 = \rho(p_T^J R)^2$, which is suppressed by $\rho$ compared to $p_c^2$. The ultracollinear mode is insensitive to the jet boundary. Figure 1(a) shows the ultracollinear mode for case (a) in Eq. (25), responsible for narrow energetic radiations inside the jet. Because $z_{\rm max}$ ($z_{\rm min}$) is of $\mathcal{O}(1)$, csoft and soft radiations do not contribute.

In case (b) with $y_c \ll 1$ in Eq. (25), in addition to the ultracollinear mode, other modes exist due to the smallness of $y_c$. In the midrange region where $\rho \sim y_c$ ($M_J^2 \sim M_c^2 \sim y_c(p_T^J R)^2$), the csoft mode, in addition to the ultracollinear mode, can contribute to the nonzero groomed jet mass, as illustrated in Fig. 1(b). Here, the csoft momentum scales as

$$p_{\rm cs}^\mu \sim y_c p_T^J(1, R^2, R) \;\; \text{with } p_{\rm cs}^2 \sim (y_c p_T^J R)^2. \qquad (27)$$

Hence, the csoft contribution to the jet mass $M_{\rm cs}^2 = p_J^+ p_{\rm cs}^-$ is power counted as $y_c(p_T^J R)^2$ and is comparable to the jet mass squared in this region. The csoft mode is the scaled-down version of the collinear mode; hence, it is sensitive to the jet boundary, as shown in Fig. 1(b).

In the peak region with $\rho \ll y_c \ll 1$, we need a much narrower version of the csoft mode to produce a much smaller nonzero jet mass. We call this mode the ultracollinear-soft (ucsoft) mode, which scales as

$$\begin{aligned} p_{\rm ucs}^\mu &\sim y_c p_T^J\Big(1, \frac{M_J^2}{y_c p_T^{J2}}, \frac{M_J}{\sqrt{y_c} p_T^J}\Big) \\ &\sim y_c p_T^J\Big(1, R^2\frac{\rho}{y_c}, R\sqrt{\frac{\rho}{y_c}}\Big), \end{aligned} \qquad (28)$$

with $p_{\rm ucs}^2 \sim y_c M_J^2$. Therefore the ucsoft mode is much narrower than the csoft mode although the largest momentum component is of the same magnitude $y_c p_T^J$. The ucsoft mode cannot recognize the jet boundary, as can be seen in Fig. 1(b).

### 1. $\rho \ll y_c \sim \mathcal{O}(1)$

In this case, the ultracollinear mode entirely describes the groomed nonzero jet mass, but this mode is insensitive to the jet boundary. In Fig. 2(a), the phase space for the ultracollinear mode is shown. The shaded region satisfies the mMDT criterion in Eq. (18) and contributes to the nonzero groomed mass. Because the invariant mass squared for the ultracollinear mode is small, the upper bound for $M^2$ extends to infinity. Outside the shaded region, $\delta(M_J^2)$ is returned because the remaining single energetic parton is included in the "grooming" procedure. As a result, at NLO in $\alpha_s$, the groomed jet functions from the ultracollinear mode are expressed as ($k = q,\, g$)

$$\begin{aligned} \tilde{J}_k(M_J^2, \mu) &= \sum_{l=q,\bar{q},g}\Bigg[\int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z z\; \tilde{\mathcal{G}}_{l/k}(z, M_J^2) \\ &\quad + \delta(M_J^2)\Big(\int_0^{z_{\rm min}} {\rm d}z + \int_{z_{\rm max}}^1 {\rm d}z\Big)\int_0^\infty {\rm d}M^2 z\; \tilde{\mathcal{G}}_{l/k}(z, M^2)\Bigg] \\ &= \delta(M_J^2) + \frac{\alpha_s C_k}{2\pi}\Big\{-\delta(M_J^2)\Big[g_k(y_c)\ln\frac{\mu^2}{\Lambda^2} + h_k(y_c)\Big] \\ &\qquad + g_k(y_c)\Big[\frac{1}{M_J^2}\Big]_{\Lambda^2}\Big\}. \end{aligned} \qquad (29)$$

Here, $\tilde{\mathcal{G}}_{l/k}(z, M^2)$ are the generic FJFs introduced in Refs. 29, 30. Unlike the $\mathcal{G}_{l/J_k}$ that was introduced in Eq. (10), $\tilde{\mathcal{G}}_{l/k}$ is not divided by $\mathcal{J}_k$. The functions $g_k$ and $h_k$ in Eq. (29) are given in Eqs. (A4) and (A6), respectively.

Even though the collinear mode does not contribute to the groomed nonzero jet mass, it can radiate in the regions $[0, z_{\rm min})$ and $(z_{\rm max}, 1]$ inside a jet. When the zero-bin subtraction [56] is employed to avoid double counting on the phase space overlapped with the ultracollinear mode, the one-loop result for the collinear mode is IR finite, and the divergence is of an UV origin. The renormalized collinear contribution to NLO in $\alpha_s$ is given by

$$\tilde{\mathcal{C}}_k(Q^2, \mu) = 1 + \frac{\alpha_s C_k}{2\pi}\Big[g_k(y_c)\ln\frac{\mu^2}{M_c^2} + h_k(y_c) - I_k(y_c)\Big], \qquad (30)$$

where we use $Q \equiv p_T^J R$ for simplicity, and $M_c^2 = Q^2 y_c/(1+y_c)^2$. In Appendix C, we show the details of the calculation.

As a result, the groomed jet mass distribution functions $\Phi_k$ ($k = q,\, g$) in the limit $\rho \ll y_c \sim \mathcal{O}(1)$ can be factorized as

$$\Phi_k(y_c \sim \mathcal{O}(1), M_J^2 \ll Q^2) = \tilde{\mathcal{C}}_k(Q^2, \mu)\tilde{J}_k(M_J^2, \mu). \qquad (31)$$

Combining Eqs. (29) and (30), one can verify that the NLO result in Eq. (31) reproduces the full result in Eq. (24) in the limit $M_J^2 \ll Q^2$. Note that each factorized function in Eq. (31) contains a single logarithm at one loop. This implies that the dominant logarithmic corrections appear as $\sum_{n=0} c_n(\alpha_s L)^n$ in the limit $\rho \ll y_c \sim \mathcal{O}(1)$, which is a typical feature of the mMDT [2].

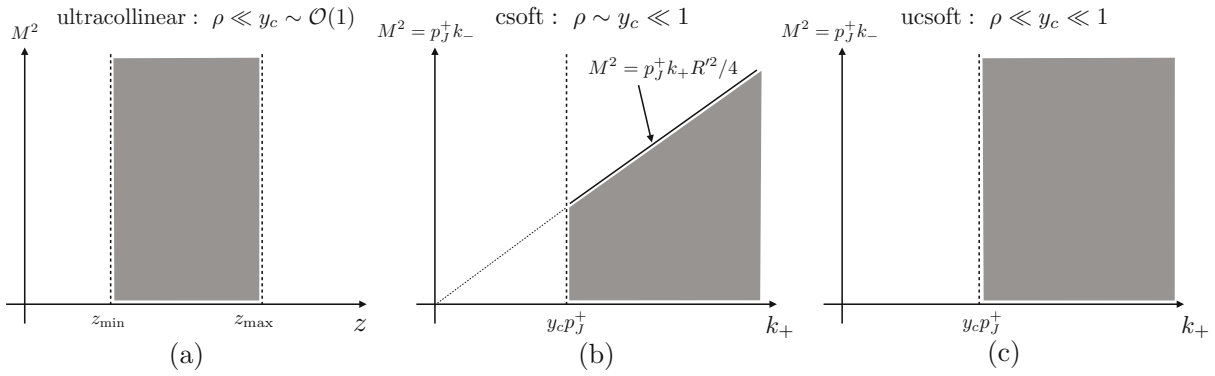


Fig. 2. Phase spaces (shaded regions) for the submodes satisfying the mMDT criterion in Eq.(18) at one loop. Diagrams (b) and (c) denote the phase spaces for the csoft and the ucsoft modes with momentum $k$. The nonzero groomed mass comes from the shaded regions, and $\delta(M_J^2)$ comes from the remaining regions according to grooming.

**2. The midrange region: $\rho \sim y_c \ll 1$**

In this region, collinear modes cannot be emitted inside the jet because the jet mass is too small, and they cannot satisfy the mMDT criterion. Therefore, the collinear contribution genuinely becomes the normalization factor $C_k = \mathcal{J}_k^{-1}$ like the contribution to the plain jet mass distribution in the limit $\rho \ll 1$, as shown in Eq. (14). The NLO results for $\mathcal{J}_k(Q^2)$ are given in Eqs. (7) and (8).

Then, as explained previously, the nonzero groomed jet mass can be described by using the ultracollinear mode and the csoft mode. Because the ultracollinear mode cannot recognize the (ungroomed) jet boundary and the mMDT criterion, the contribution yields the standard jet functions, which are given to NLO as

$$\begin{aligned} J_q(M^2,\mu) &= \delta(M^2) \\ &\quad + \frac{\alpha_s C_F}{2\pi}\Big\{\delta(M^2)\Big[\frac{3}{2}\ln\frac{\mu^2}{\Lambda^2} + \ln^2\frac{\mu^2}{\Lambda^2} + \frac{7}{2} - \frac{\pi^2}{2}\Big] \\ &\qquad - \Big[\Big(2\ln\frac{\mu^2}{M^2} + \frac{3}{2}\Big)\frac{1}{M^2}\Big]_{\Lambda^2}\Big\}, \\ J_g(M^2,\mu) &= \delta(M^2) \\ &\quad + \frac{\alpha_s}{2\pi}\Big\{\delta(M^2)\Big[C_A\Big(\ln\frac{\mu^2}{\Lambda^2} + \frac{67}{18} - \frac{\pi^2}{2}\Big) \\ &\qquad - \frac{10}{9}T_R n_f + \frac{\beta_0}{2}\ln\frac{\mu^2}{\Lambda^2}\Big] \\ &\qquad - \Big[\Big(\frac{\beta_0}{2} + 2C_A\ln\frac{\mu^2}{M^2}\Big)\frac{1}{M^2}\Big]_{\Lambda^2}\Big\}. \end{aligned} \tag{32}$$

The decoupled csoft gluons from the collinear fields form the csoft Wilson lines $Y_{n(\overline{n}),\mathrm{cs}}$. Similar to Eq. (17), the contribution to the groomed jet mass (with an energetic quark) can be expressed as

$$\begin{aligned} S^I_{k=q}(M^2_{cs} = p_J^+\ell_-,\mu) &= \frac{1}{p_J^+ N_c} \\ &\times \mathrm{Tr}\,\langle 0|Y^\dagger_{\overline{n},\mathrm{cs}} Y_{n,\mathrm{cs}}\delta(\ell_- + \Theta_{\mathrm{cs}} i\partial_-) Y^\dagger_{n,\mathrm{cs}} Y_{\overline{n},\mathrm{cs}}|0\rangle, \end{aligned} \tag{33}$$

and $S_g^I$ is expressed in terms of the Wilson lines in the adjoint representation. Here $\Theta_{\mathrm{cs}}$ represents the mMDT criterion in Eq. (18), which the csoft mode should pass. The phase space for the csoft mode to pass the criterion is illustrated as a shaded region in Fig. 2(b). In the shaded region, the csoft function $S_k^I$ contributes to the nonzero groomed jet mass, while the contributions from the remaining region and the virtual correction yield the part proportional to $\delta(M^2)$.

The contribution from the shaded region in Fig. 2(b) involves the IR divergence as $M$ goes to zero. This IR divergence is proportional to $\delta(M^2)$ in the $\Lambda$-distribution defined in Eq. (23). Then, the IR divergence in the real emission, appearing in the part with $\delta(M^2)$, is cancelled by the virtual contribution. As a result, only the UV divergence remains, and we find that the renormalized csoft functions to NLO are given as

$$\begin{aligned} S_k^I(M^2,\mu) = \delta(M^2) &+ \frac{\alpha_s C_k}{2\pi}\Big\{-\delta(M^2)\Big(\frac{1}{2}\ln\frac{\mu^2}{y_c^2 Q^2} - \frac{\pi^2}{12}\Big) \\ &+ \frac{2}{M^2}\Theta(M^2 - M_c^2)\ln\frac{\mu^2 Q^2}{(M^2)^2} \\ &+ \Big[\frac{2}{M^2}\Theta(M_c^2 - M^2)\ln\frac{\mu^2}{y_c M^2}\Big]_{M_c^2}\Big\}, \end{aligned} \tag{34}$$

where $M_c^2 = y_c Q^2$, and we set the upper bound to $\Lambda^2 = M_c^2$ in the $\Lambda$-distribution. The results with the mMDT, along with those with soft drop in Appendix D.1, are new.

Finally the groomed jet mass distribution functions $\Phi_k$

in the limit $\rho \sim y_c \ll 1$ can be factorized as

$$\Phi_k(y_c \ll 1, M_J^2 \sim y_c Q^2) = \mathcal{C}_k(Q^2,\mu) \times \int_0^{M_J^2} \mathrm{d}M^2 J_k(M^2,\mu) S_k^I(M_J^2 - M^2,\mu). \quad (35)$$

The factorization structure is the same as the factorized expression for the plain jet mass in the limit $\rho \ll 1$. [See Eq. (14).] What is different from the plain jet mass distribution is that the csoft function $S_k^I$ in Eq. (35) is affected by the mMDT criterion. Note that both the groomed and the ungroomed jet mass distributions are scale invariant. Therefore, the renormalization behavior of $S_k^I$ is the same in both cases. This will be explained in detail in Sec. V.

Expanding all the terms to NLO in the factorized parts, the groomed jet mass distribution functions $\Phi_k^{(1)}(y_c, M^2)$ at fixed order in $\alpha_s$ are given as

$$\Phi_q^{(1)}(y_c, M_J^2) = \frac{\alpha_s C_F}{2\pi}\Big\{\delta(M_J^2)\Big(-\ln^2 y_c - \frac{3}{2}\ln y_c - 3 + \frac{\pi^2}{3}\Big) + \frac{1}{M_J^2}\Big(2\ln\frac{Q^2}{M_J^2} - \frac{3}{2}\Big)\Theta(M_J^2 - M_c^2) - \Big(2\ln y_c + \frac{3}{2}\Big)\Big[\frac{\Theta(M_c^2 - M_J^2)}{M^2}\Big]_{M_c^2}\Big\}, \quad (36)$$

$$\Phi_g^{(1)}(y_c, M_J^2) = \frac{\alpha_s}{2\pi}\Big\{\delta(M_J^2)\Big[C_A\Big(-\ln^2 y_c - \frac{67}{18} + \frac{\pi^2}{3}\Big) - \frac{\beta_0}{2}\ln y_c + T_R n_f \frac{13}{9}\Big] + \frac{2C_A}{M_J^2}\ln\frac{\mu^2 Q^2}{(M_J^2)^2}\Theta(M_J^2 - M_c^2) - \Big(\frac{\beta_0}{2} + 2C_A \ln y\Big)\Big[\frac{1}{M_J^2}\Big]_{M_c^2}\Theta(M_c^2 - M_J^2)\Big\}. \quad (37)$$

These are consistent with the results in Eq. (24) in the limit $y_c \ll \mathcal{O}(1)$.

### 3. The peak region: $\rho \ll y_c \ll 1$

In the peak region, distinct contributions from the collinear modes with $p_c^2 \sim (p_T^J R)^2$, the csoft modes with $p_{\rm cs}^2 \sim y_c^2 (p_T^J R)^2$, the ultracollinear modes with $p_{\rm uc}^2 \sim M_J^2$ and finally the ucsoft modes with $p_{\rm ucs}^2 \sim y_c M_J^2$ participate in the groomed jet mass distribution $\Phi_k$. The collinear mode cannot radiate inside a jet because $\rho \ll 1$, and the csoft mode cannot satisfy $\rho \ll y_c$. Therefore, none of these modes contributes to nonzero jet mass; hence, these modes contribute only to the normalization. On the other hand, the ultracollinear and the ucsoft modes can contribute to the nonzero groomed jet mass.

The factorized groomed jet mass distribution functions can be written as ($k = q,\, g$)

$$\Phi_k(y_c \ll 1, M_J^2 \ll y_c Q^2) = \mathcal{C}_k(Q^2,\mu)\mathcal{S}_k^{II}(y_c^2 Q^2,\mu) \times \int_0^{M_J^2} \mathrm{d}M^2 J_k(M^2,\mu) U_k(M_J^2 - M^2,\mu). \quad (38)$$

The collinear functions $\mathcal{C}_k$ are the inverses of the integrated jet functions $\mathcal{J}_k^{-1}(Q^2,\mu)$ as in Eqs. (14) and (35). However, $\mathcal{S}_k^{II}$ are the csoft contributions in this region, different from $S_k^I$ in Eq. (35). $J_k$ are the standard jet functions for the ultracollinear mode, that also appear in Eqs. (14) and (35). Finally $U_k$ are the ucsoft contributions to the jet mass distributions.

The detailed calculation of $\mathcal{S}_k^{II}$ is shown in Appendix C. After the zero-bin subtraction, the NLO results are free of the IR divergence, and the renormalized results are given as

$$\mathcal{S}_k^{II}(y_c^2 Q^2,\mu) = 1 + \frac{\alpha_s C_k}{2\pi}\Big(\frac{1}{2}\ln^2\frac{\mu^2}{y_c^2 Q^2} - \frac{\pi^2}{12}\Big). \quad (39)$$

The ucsoft functions $U_k$ in Eq. (38) can be defined in a similar way to $S_k^I$ in Eq. (33), except that the Wilson lines are replaced by those with the ucsoft gauge fields and the conditional function $\Theta_{\rm cs}$ is replaced by $\Theta_{\rm ucs}$ due to the different scaling of the ucsoft mode. As a result, the quark ucsoft function is defined as

$$U_q(M^2 = p_J^+ \ell_-,\mu) = \frac{1}{p_J^+ N_c} \times \mathrm{Tr}\,\langle 0|Y_{\overline{n},\rm ucs}^\dagger Y_{n,\rm ucs}\delta(\ell_- + \Theta_{\rm ucs} i\partial_-) Y_{n,\rm ucs}^\dagger Y_{\overline{n},\rm ucs}|0\rangle. \quad (40)$$

Here, $\Theta_{\rm ucs} = \Theta(-i\partial_+ - y_c p_J^+)$; hence, the ucsoft contribution to the nonzero mass comes from the shaded region in Fig. 2(c). The ucsoft mode scales as Eq. (28) and cannot recognize the jet boundary. The gluon ucsoft function $U_g$ can be similarly defined in terms of the Wilson lines in the adjoint representation.

Similar to computing $S_k^I$, we can calculate the NLO contributions to $U_k$ by employing the $\Lambda$-distribution. Then the IR divergences are cancelled by the virtual contributions, and the renormalized results to NLO are given as

$$U_k(M^2,\mu) = \delta(M^2)\Big[1 - \frac{\alpha_s C_k}{\pi}\Big(\frac{1}{2}\ln^2\frac{\mu^2}{y_c\Lambda^2} - \frac{\pi^2}{12}\Big)\Big] + \frac{\alpha_s C_k}{\pi}\Big[\frac{1}{M^2}\ln\frac{\mu^2}{y_c M^2}\Big]_{\Lambda^2}. \quad (41)$$

By expanding the factorized functions in Eq. (38) to order $\alpha_s$, we verify that the NLO results are consistent with the results in Eq. (24) [or Eqs. (36) and (37)] after taking the limit $\rho \ll y_c \ll 1$. The ucsoft function is also computed using a soft drop, and it is given in Appendix D.2.

## V. RESUMMED GROOMED JET MASS DISTRIBUTIONS

We perform the resummation of the large logarithms arising from small $y_c$ and $\rho$ in the groomed jet mass distributions. Though we will fix $y_c = 0.1$ in the numerical analysis, here, we regard $y_c$ as a very small number in order to investigate the resummed effects of the large logarithms in $y_c$. We have two distinctive small jet-mass regions: the midrange region ($\rho \sim y_c \ll 1$) and the peak region ($\rho \ll y_c \ll 1$). These two regions have different factorized structures, hence different resummation effects.

The resummation of large logarithms can be achieved by solving the RG equation of the factorized parts. We first factorize the groomed jet mass distributions and find the appropriate scales at which the logarithms of the factorized functions are minimized. Then, each factorized function is evolved by using the RG equation to a common factorization scale $\mu_f$. Combining all the RG evolutions of each factorized function yields the resummation of the large logarithms.

The groomed jet mass distribution functions in the midrange and the peak regions involve double logarithms at NLO, so the resummed result to leading logarithmic (LL) accuracy, $\sum_{n=0} a_k(\alpha_s L^2)^n \sim \exp(L f_0(\alpha_s L))$, is estimated to be larger than the $\mathcal{O}(1)$ contributions. Here, $L$ represents the large logarithms in small $\rho$ or $y_c$. Hence, in order to include $\mathcal{O}(1)$ contributions, we resum the large logarithms up to NLL accuracy, which is schematically given as $\sum_{n=0} b_k(\alpha_s L)^n \sim \exp(f_1(\alpha_s L)) \sim \mathcal{O}(1)$.

From now on, we use the dimensionless jet mass variable $\rho = M_{(J)}^2/Q^2$ to consider the dimensionless functions for the jet mass distribution and its factorized functions.[3] The relation between the dimensionless functions $\bar{f}(\rho)$ and the dimensionful functions $f(M^2)$ is given by $\bar{f}(\rho) = Q^2 f(M^2)$, where $f = \{\Phi_k, \tilde{J}_k, J_k, S_k^I, U_k\}$. With the dimensionless functions, we can express the $\Lambda$-distributions in the dimensional functions in terms of the standard plus distributions and $\delta(\rho)$. The standard plus distribution is defined as

$$\int_0^x \mathrm{d}\rho [g(\rho)]_+ h(\rho) = \int_0^x \mathrm{d}\rho g(\rho) h(\rho) - h(0) \int_0^1 \mathrm{d}\rho g(\rho), \tag{42}$$

where $h(\rho)$ is an arbitrary function that is smooth at $\rho = 0$. The details of the conversion are shown in Appendix B.

[3] Throughout this paper, the dimensionless jet mass variable $\rho$ indicates $M_J^2/Q^2$ in most cases. However, it is sometimes used for expressing a partial jet mass squared over $Q^2$ (*i.e.*, $M^2/Q^2$) when we consider dimensionless factorized functions such as $\bar{J}_k(\rho)$ and $\bar{U}_k(\rho)$.

### 1. Midrange region: $\rho \sim y_c \ll 1$

The factorization theorem in the midrange region is given in Eq. (35), and the factorized functions satisfy the following RG equations:

$$\frac{d}{d\ln\mu}\mathcal{C}_k = \gamma_{\mathcal{C}}^k \mathcal{C}_k, \quad \frac{d}{d\ln\mu}\bar{f}_k(\rho) = \int_0^\rho \mathrm{d}\rho' \gamma_f^k(\rho') \bar{f}_k(\rho-\rho'), \tag{43}$$

where $k = q,\ g$ and $\bar{f}_k = \bar{J}_k,\ \bar{S}_k^I$. The anomalous dimensions in general can be expressed as

$$\gamma_{\mathcal{C}}^k = A_c \Gamma_C^k \ln\frac{\mu^2}{Q^2} + \hat{\gamma}_c^k, \tag{44}$$

$$\gamma_J^k(\rho) = \delta(M^2)\Big(A_j \Gamma_C^k \ln\frac{\mu^2}{Q^2} + \hat{\gamma}_j^k\Big) - \kappa_j A_j \Gamma_C^k \Big[\frac{1}{\rho}\Big]_+, \tag{45}$$

$$\gamma_{S^I}^k(\rho) = \delta(\rho)\Big(A_s \Gamma_C^k \ln\frac{\mu^2}{Q^2} + \hat{\gamma}_{s1}^k\Big) - \kappa_s A_s \Gamma_C^k \Big[\frac{1}{\rho}\Big]_+, \tag{46}$$

where $\Gamma_C^k = \sum_{n=0} \Gamma_n^k (\alpha_s/4\pi)^{n+1}$ are the cusp anomalous dimensions [57,58], and the first two coefficients are given by

$$\Gamma_0^k = 4C_k, \quad \Gamma_1^k = 4C_k\left[\Big(\frac{67}{9} - \frac{\pi^2}{3}\Big)C_A - \frac{10}{9}n_f\right], \tag{47}$$

with $C_q = C_F$ and $C_g = C_A$. From the NLO results in Eqs. (7), (8), (32), and (34), we extract the set of coefficients $\{A_c, A_j, A_s, \kappa_j, \kappa_s\} = \{-1, 2, -1, 1, 2\}$. Furthermore, the noncusp anomalous dimensions to NLL accuracy are given by $\hat{\gamma}_c^q = -3\alpha_s C_F/(2\pi)$, $\hat{\gamma}_c^g = -\alpha_s\beta_0/(2\pi)$, $\hat{\gamma}_j^k = -\hat{\gamma}_c^k$, and $\hat{\gamma}_{s1}^k = 0$.

We find that the factorization structure for the groomed jet mass distribution is the same as that for the plain jet mass distribution in the limit $\rho \ll 1$ (See Ref. 19). In this limit, the grooming effects (represented by $y_c$) appear only in the csoft function. However, the RG behavior for the csoft function should be the same as that for the plain jet mass because both jet mass distributions are scale invariant and the collinear function $\mathcal{C}_k$ and the jet function $J_k$ are the same in both cases.

The RG equations can be solved by following the conventional methods using the Laplace transform [59,60], and the resummed result at NLL accuracy can be written as

$$\begin{aligned}\bar{\Phi}_k(\rho \sim y_c) &= \Phi_k(M_J^2)\cdot Q^2 \\ &= \exp[\mathcal{M}_k^I(\mu_c, \mu_{uc}, \mu_{cs})]\ \mathcal{C}_k(Q, \mu_c) \\ &\quad \times \hat{J}_k\Big[\ln\frac{\mu_{uc}^2}{Q^2} - \partial_{\eta_1}\Big]\hat{S}_k^I\Big[\ln\frac{\mu_{cs}^2}{Q^2} - 2\partial_{\eta_1}\Big]\frac{e^{-\gamma_E\eta_1}}{\Gamma(\eta_1)}\rho^{-1+\eta_1},\end{aligned} \tag{48}$$

where $\bar{\Phi}_k(\rho)$ is the dimensionless jet mass distribution with $\rho = M_J^2/Q^2$ and $Q = p_T^J R$. Moreover, $\hat{J}_k$ and $\hat{S}_k$ in

Eq. (48) are the Laplace transforms of the dimensionless functions $\bar{J}_k(\rho)$ and $\bar{S}_k^I(\rho)$, which are given to NLO as

$$\hat{J}_q[L] = 1 + \frac{\alpha_s C_F}{2\pi}\Big(\frac{7}{2} - \frac{\pi^2}{3} + \frac{3}{2}L + L^2\Big), \tag{49}$$

$$\hat{J}_g[L] = 1 + \frac{\alpha_s C_A}{2\pi}\Big(\frac{67}{18} - \frac{5n_f}{9C_A} - \frac{\pi^2}{3} + \frac{\beta_0}{2C_A}L + L^2\Big), \tag{50}$$

$$\hat{S}_k^I[L] = 1 + \frac{\alpha_s C_k}{2\pi}\Big(-\frac{1}{2}L^2 - \frac{\pi^2}{4}\Big). \tag{51}$$

In Eq. (48), $\eta_1 = 2a[\Gamma_C^k](\mu_{\rm uc}, \mu_{\rm cs})$ and is positive for $\mu_{\rm uc} > \mu_{\rm cs}$. Because $\Phi_k$ is scale invariant, the factorization scale $\mu_f$ does not appear in Eq. (48).

The exponent $\mathcal{M}_k^I$ to NLL accuracy can be written as

$$\begin{aligned}\mathcal{M}_k^I(\mu_{\rm c}, \mu_{\rm uc}, \mu_{\rm cs}) &= -2S_\Gamma^k(\mu_{\rm uc}, \mu_{\rm c}) - 2S_\Gamma^k(\mu_{\rm uc}, \mu_{\rm cs}) \\ &\quad + \ln\frac{\mu_{\rm uc}^2}{Q^2}\Big(a[\Gamma_C^k](\mu_{\rm c}, \mu_{\rm uc}) + a[\Gamma_C^k](\mu_{\rm cs}, \mu_{\rm uc})\Big) \\ &\quad + a[\hat{\gamma}_j^k](\mu_{\rm c}, \mu_{\rm uc}).\end{aligned} \tag{52}$$

Here $S_\Gamma^k$ and $a[f]$ are defined as

$$\begin{aligned}S_\Gamma^k(\mu_1, \mu_2) &= \int_{\alpha_2}^{\alpha_1}\frac{d\alpha_s}{b(\alpha_s)}\Gamma_C^k(\alpha_s)\int_{\alpha_1}^{\alpha_s}\frac{d\alpha_s'}{b(\alpha_s')}, \\ a[f](\mu_1, \mu_2) &= \int_{\alpha_2}^{\alpha_1}\frac{d\alpha_s}{b(\alpha_s)}f(\alpha_s),\end{aligned} \tag{53}$$

where $\alpha_{1,2} \equiv \alpha_s(\mu_{1,2})$ and $b(\alpha_s) = d\alpha_s/d\ln\mu$ is the QCD beta function.

### 2. Peak region: $\rho \ll y_c \ll 1$

The groomed jet mass distribution can be factorized in terms of the collinear ($\mathcal{C}_k$), the csoft ($\mathcal{S}_k^{II}$), the ultracollinear ($J_k$), and the ucsoft ($U_k$) functions, as shown in Eq. (38). Here, the collinear functions $\mathcal{C}_k$ and the jet functions $J_k$ for the ultracollinear modes are the same as those in the midrange region. The RG equations for $\mathcal{S}_k^{II}$ and the dimensionless function $\bar{U}_k$ are given by

$$\begin{aligned}\frac{d}{d\ln\mu}\mathcal{S}_k^{II} &= \gamma_{\mathcal{S}^{II}}^k\mathcal{S}_k^{II}, \\ \frac{d}{d\ln\mu}\bar{U}_k(\rho) &= \int_0^\rho {\rm d}\rho'\gamma_U^k(\rho')\bar{U}_k(\rho - \rho').\end{aligned} \tag{54}$$

Moreover, the anomalous dimensions can be written as

$$\gamma_{\mathcal{S}^{II}}^k = A_{s2}\Gamma_C^k\ln\frac{\mu^2}{y_c^2Q^2} + \hat{\gamma}_{s2}^k, \tag{55}$$

$$\gamma_U^k(\rho) = \delta(\rho)\Big(A_u\Gamma_C^k\ln\frac{\mu^2}{y_cQ^2} + \hat{\gamma}_u^k\Big) - \kappa_u A_u\Gamma_C^k\Big[\frac{1}{\rho}\Big]_+. \tag{56}$$

The coefficients $\{A_{s2}, A_u, \kappa_u\} = \{1, -2, 1\}$ are obtained from the NLO results in Eqs. (39) and (41), and the noncusp anomalous dimensions to NLL accuracy are $\hat{\gamma}_{s2}^k = \hat{\gamma}_u^k = 0$. The scale invariance of $\Phi_k(\rho \ll y_c)$ is guaranteed by the following relations:

$$A_c + A_{s2} + A_j + A_u = 0, \quad \kappa_j A_j + \kappa_u A_u = 0, \quad \hat{\gamma}_c^k + \hat{\gamma}_j^k = 0. \tag{57}$$

By evolving the factorized functions from their own scales to the factorization scale $\mu_f$ from Eq. (54), we can write the resummed results as

$$\begin{aligned}\bar{\Phi}_k(\rho \ll y_c) &= \exp[\mathcal{M}_k^{II}(\mu_{\rm c}, \mu_{\rm cs}, \mu_{\rm uc}, \mu_{\rm ucs})] \\ &\quad \times \mathcal{C}_k(Q, \mu_{\rm c})\mathcal{S}_k^{II}(y_cQ, \mu_{\rm cs})\hat{J}_k\Big[\ln\frac{\mu_{\rm uc}^2}{Q^2} - \partial_{\eta_2}\Big] \\ &\quad \times \hat{U}_k\Big[\ln\frac{\mu_{\rm ucs}^2}{Q^2} - \partial_{\eta_2}\Big]\frac{e^{-\gamma_E\eta_2}}{\Gamma(\eta_2)}\rho^{-1+\eta_2},\end{aligned} \tag{58}$$

where $\eta_2 = 2a[\Gamma_C^k](\mu_{\rm uc}, \mu_{\rm ucs})$. Also, $\hat{U}_k$ are the Laplace transforms of the dimensionless ucsoft functions $\bar{U}_k(\rho)$, which are given to NLO as

$$\hat{U}_k[L] = 1 - \frac{\alpha_s C_k}{2\pi}L^2. \tag{59}$$

The exponent $\mathcal{M}_k^{II}$ can be written as

$$\begin{aligned}\mathcal{M}_k^{II}(\mu_{\rm c}, \mu_{\rm cs}, \mu_{\rm uc}, \mu_{\rm ucs}) &= 2S_\Gamma^k(\mu_{\rm c}, \mu_{\rm cs}) - 4S_\Gamma^k(\mu_{\rm uc}, \mu_{\rm ucs}) \\ &\quad + \ln\frac{\mu_{\rm c}^2}{Q^2}a[\Gamma_C^k](\mu_{\rm c}, \mu_{\rm cs}) - 2\ln\frac{\mu_{\rm uc}^2}{Q^2}a[\Gamma_C^k](\mu_{\rm uc}, \mu_{\rm ucs}) \\ &\quad + 2\ln y_c a[\Gamma_C^k](\mu_{\rm cs}, \mu_{\rm ucs}) + a[\hat{\gamma}_j^k](\mu_{\rm c}, \mu_{\rm uc}).\end{aligned} \tag{60}$$

### 3. Nonglobal logarithms

NGLs [54,55] arise when gluons are radiated across the jet boundary and contribute to a jet observable with the phase space constrained by the boundary. Although the leading NGLs begin to appear at two loop, the perturbative series is schematically given as $\sum_{n=2} b_{\rm NG}^n(\alpha_s L_{\rm NGL})^n$ and hence, contributes to NLL accuracy. Especially when a large energy difference exists between the gluons across the boundary, large NGLs appear. In the effective theory approach, when multiple modes are involved to resolve the jet boundary and a large energy hierarchy exists among them, large NGLs appear.

The groomed jet mass in the midrange region ($\rho \sim y_c \ll 1$) is a NGL observable because the collinear and the csoft modes can resolve the jet boundary, and the csoft mode can contribute to the groomed jet mass. In the peak region ($\rho \ll y_c \ll 1$), although the NGLs can be generated by the collinear and the csoft modes, they do not affect the jet mass directly because the collinear

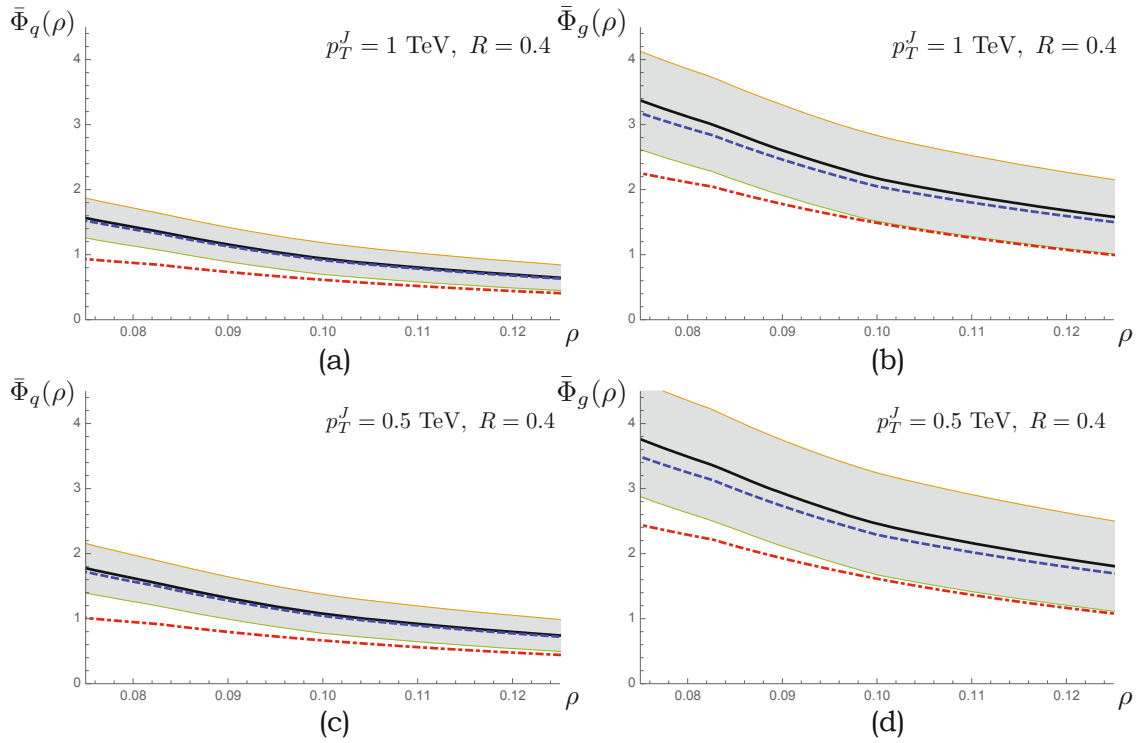


Fig. 3. (Color online) Groomed jet mass distributions multiplied by $\rho$ for (a, c) a quark-initiated jet and (b, d) a gluon-initiated jet with $(p_T^J, R) = (1\ \text{TeV}, 0.4)$ and $(p_T^J, R) = (0.5\ \text{TeV}, 0.4)$ in the midrange region ($\rho \sim y_c \ll 1$). Here, $\rho = M_J^2/(p_T^J R)^2$. The black thick lines denote the NLL resummed results with the fixed NLO corrections ($\text{NLL}_\text{G} + \text{NLO}$). The blue dashed lines denote the resummed results including the NGLs ($\text{NLL}_{\text{G+NG}} + \text{NLO}$). The red dot-dashed lines are the results at NLO without resummation. Gray bands at $\text{NLL}_\text{G} + \text{NLO}$ show the uncertainties under the scale variations from $\mu_{c,uc,cs} = 2\mu^0_{\text{c,uc,cs}}$ to $\mu^0_{\text{c,uc,cs}}/2$.

and the csoft modes contribute only to the normalization of the jet mass distribution.

For the plain jet mass distributions, the contribution of the NGLs is sizable around the peak region, but decreases rapidly away from the peak. However, we expect the groomed jet mass not to be affected substantially by the NGLs. The groomed jet mass in the peak region is not a NGL observable at all, and the NGL contribution in the midrange region is quite suppressed because that region is far away from the peak.

Remarkably the resummed result of the leading NGLs for a narrow isolated jet will take the same form as that for a hemisphere jet mass because the generating mechanisms of the NGLs are similar [5,8]. Therefore, in order to estimate the NGL contribution to the groomed jet mass, we might be able to use the resummed formula of the leading NGLs in the large $N_c$ limit for the hemisphere jet mass [54], which can be written as

$$\Delta_{\text{NG}}^{k=q,g}(\mu_1,\mu_2) = \exp\left(-C_A C_k \frac{\pi^2}{3}\left(\frac{1+(at)^2}{1+(bt)^c}\right)t^2\right), \quad (61)$$

where

$$t = \frac{1}{\beta_0}\ln\frac{\alpha_s(\mu_2)}{\alpha_s(\mu_1)} \sim -\frac{1}{\beta_0}\ln\left(1-\frac{\beta_0}{4\pi}\alpha_s(\mu_1)\ln\frac{\mu_1^2}{\mu_2^2}\right). \quad (62)$$

Here, the fit parameters from the Monte Carlo simulation are given by $a = 0.85C_A$, $b = 0.86C_A$, and $c = 1.33$ [54]. In our analysis, we set $(\mu_1,\mu_2) = (\mu_\text{c}, \mu_\text{cs}^I)$ for the midrange region and $(\mu_1,\mu_2) = (\mu_\text{c}, \mu_\text{cs}^{II})$ for the peak region. For the NLL-resummed results (with the fixed NLO), the contributions of the NGLs in Eq. (61) is multiplicative in Eq. (48) or Eq. (58).

## VI. NUMERICAL RESULTS

We present the numerical analysis of the groomed jet mass distributions $\bar{\Phi}_k(\rho)$. In fact, they are not physical observables, and for phenomenology, the ratio of the scattering cross sections in Eqs. (3) and (1) with respect to the jet mass should be considered by summing over all the contributions combined with the parton distribution functions. For example, a phenomenological analysis in the peak region is reported in Ref. 27. However, our main focus is the theoretical issue on how to implement the grooming method to the jet mass distribution and how the resummation on $\rho$ and $y_c$ affects the jet mass distribution over a wide range of $\rho$. Therefore, the purpose of the numerical analysis is to offer a theoretical under-

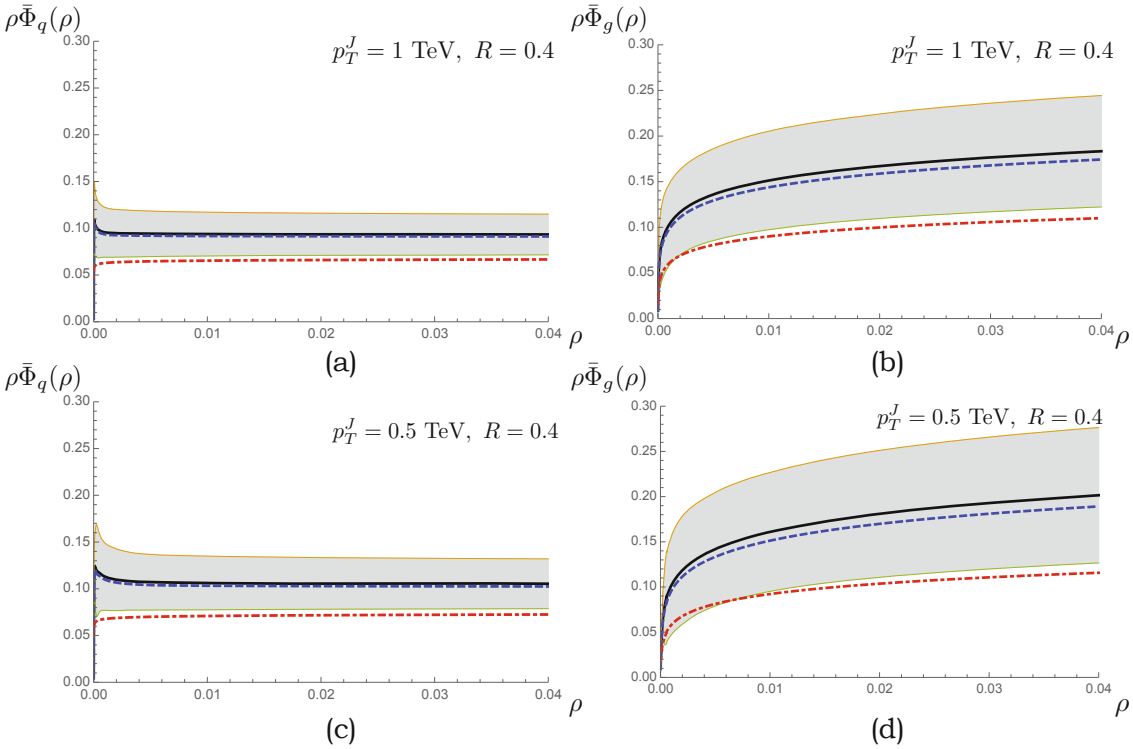


Fig. 4. (Color online) Groomed jet mass distributions multiplied by $\rho$ in the peak region ($\rho \ll y_c$): (a, c) are for the quark-initiated jet and (b, d) for the gluon-initiated jet. The black thick lines denote the NLL resummed results with the fixed NLO corrections ($\mathrm{NLL_G} + \mathrm{NLO}$), and the blue dashed lines denote the resummed results including NGLs ($\mathrm{NLL_{G+NG}} + \mathrm{NLO}$). The red dot-dashed lines are the results resumming $\ln\rho$ to LL accuracy. Gray bands for the results at $\mathrm{NLL_G} + \mathrm{NLO}$ are the uncertainties under the scale variations from $\mu_{\mathrm{c,cs,uc,ucs}} = 2\mu^0_{\mathrm{c,cs,uc,ucs}}$ to $\mu^0_{\mathrm{c,cs,uc,ucs}}/2$.

standing of how the groomed jet mass distributions for the quark- and gluon-initiated jets behave with grooming and how the results in different regions are affected by the resummation. A complete phenomenological analysis is beyond the scope of this paper and will be performed in future work.

For numerical analysis, we set $y_c = 0.1$. First of all, let us consider the behavior of the groomed jet mass distribution in the new region, which we call the midrange region ($\rho \sim y_c \ll 1$) in Fig. 3. Here, black thick lines show the results of resumming large logarithms of small $\rho$ and $y_c$ to NLL accuracy with the fixed NLO corrections ($\mathrm{NLL_G} + \mathrm{NLO}$), where $\mathrm{NLL_G}$ means the resummation without NGLs. These results can be directly obtained from Eq. (48). We set the default collinear, ultracollinear, and csoft scales as $(\mu_{\mathrm{c}}^0, \mu_{\mathrm{uc}}^0, \mu_{\mathrm{cs}}^{I,0}) = (Q, \sqrt{\rho}Q, \sqrt{\rho y_c}Q)$. Though the resummed result is independent of $\mu_f$, arbitrariness exists in setting the default scale for each factorized function. The uncertainties when each scale varies from $2\mu_i^0$ to $\mu_i^0/2$ separately are shown as gray bands in Fig. 3.

Compared to the fixed NLO results without resummation (red dot-dashed lines in Fig. 3), the resummed results are significantly enhanced by 50−80%. Because the jet mass distribution for a gluon-initiated jet is broad and decreases slowly away from the peak, the distribution for the gluon-initiated jet is dominant over the quark-initiated jet in this region. Note that the contributions of the NGLs (blue dashed lines in Fig. 3) are small even though the jet mass in this region is a nonglobal variable.

In Fig. 4, we show the groomed jet mass distributions (multiplied by $\rho$) in the peak region ($\rho \ll y_c \ll 1$). Here, thick black lines are our default results to the accuracy of $\mathrm{NLL_G} + \mathrm{NLO}$, which were obtained by using Eq. (58). The default scales for the factorized functions are chosen as $(\mu_{\mathrm{c}}^0, \mu_{\mathrm{cs}}^{II,0}, \mu_{\mathrm{uc}}^0, \mu_{\mathrm{ucs}}^0) = (Q, y_cQ, \sqrt{\rho}Q, \sqrt{\rho y_c}Q)$. We estimate the uncertainties by varying the scales from $2\mu_i^0$ to $\mu_i^0/2$ separately, and they are shown as gray bands in Fig. 4. To avoid the Landau pole as $\rho$ goes to zero, we introduce a small, fixed point, $\rho_0$. Then, in the region $\rho < \rho_0$, we freeze the ucsoft scale at a value slightly above the Landau pole. This is implemented by using the scale profile as

$$\mu_{\mathrm{ucs}}^{\mathrm{pf}} = \begin{cases} Q\sqrt{y_c\rho} & \text{if } \rho \geq \rho_0, \\ \mu_{\mathrm{min}} + aQ\sqrt{y_c\rho^3} & \text{if } \rho < \rho_0, \end{cases} \tag{63}$$

where we set $\mu_{\mathrm{min}} = 0.5$ GeV. Also, $\rho_0$ and $a$ are determined for $\mu_{\mathrm{ucs}}^{\mathrm{pf}}$ to be smooth and continuous at $\rho_0$. Accordingly, the ultracollinear scale profile is given as $\mu_{\mathrm{uc}}^{\mathrm{pf}} = \mu_{\mathrm{ucs}}^{\mathrm{pf}}/\sqrt{y_c}$.

The NGL contributions to the resummed results (blue dashed lines in Fig. 4) come from the scale deviation

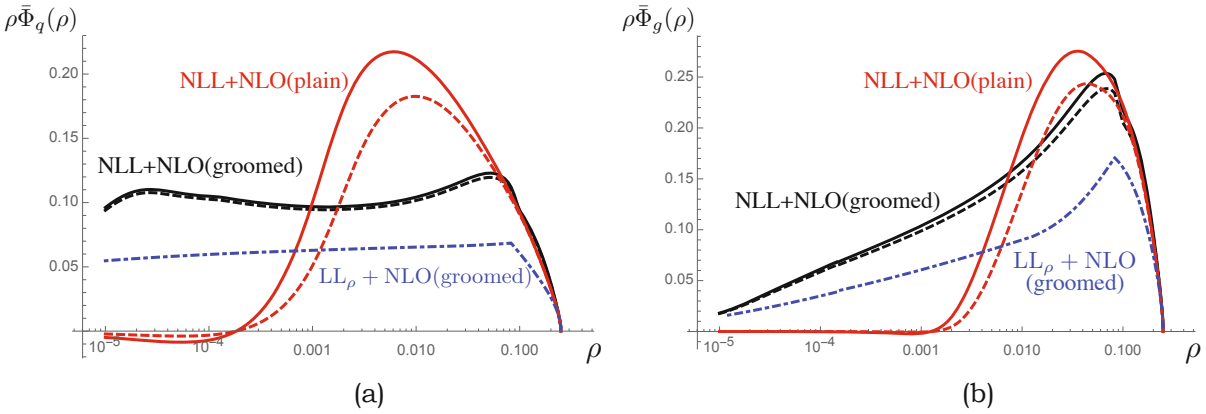


Fig. 5. (Color online) Groomed jet mass distributions multiplied by $\rho$ in the full range of $\rho$ for a quark-initiated jet [Fig. 5(a)] and a gluon-initiated jet [Fig. 5(b)] with $p_T^J = 1$ TeV and $R = 0.4$. Here black solid (dashed) lines are the groomed jet mass distributions at the accuracy of $\mathrm{NLL_G} + \mathrm{NLO}$ ($\mathrm{NLL_{G+NG}} + \mathrm{NLO}$). Red solid (dashed) lines are the plain jet mass distributions at the accuracy of $\mathrm{NLL_G} + \mathrm{NLO}$ ($\mathrm{NLL_{G+NG}} + \mathrm{NLO}$). Blue dot-dashed lines denote the groomed jet mass distributions with resummation of only $\ln\rho$ at LL accuracy.

between $\mu_c$ and $\mu_{cs}^{II}$; hence, the NGLs in the peak region take the form of $\ln y_c$ and affect the normalization of the distributions. As in the case of the midrange region, the effects are tiny and especially negligible for a quark-initiated jet.

The red dot-dashed lines in Fig. 4 are the results of resumming $\ln\rho$ only to an accuracy of LL+NLO,[4] which are based on the factorized results in Eq. (31) reflecting the limit $\rho \ll y_c \sim \mathcal{O}(1)$. Comparing them to the default results (thick black lines in Fig. 4), we see large deviations, which indicate that the resummation of $\ln y_c$ gives rise to a significant enhancement in the peak region. The enhancement from resumming on $\ln y_c$ is persistent from the midrange region to the peak region.

In Fig. 5, we finally show the groomed jet mass distributions (multiplied by $\rho$) over the full range of $\rho$. Here, black solid (dashed) lines are the fully resummed results for small $\rho$ and $y_c$ to NLL accuracy without (with) NGL effects, and they are illustrated by combining the tail ($\rho \gg y_c$), the midrange ($\rho \sim y_c$) and the peak regions ($\rho \ll y_c$). Here, the distributions for the tail region are the same as the plain jet mass distributions without resummation.

In combining the groomed jet mass distributions in three different regions, we interpolate around $\rho \sim 0.08$ ($\rho \sim 0.12$) for a smooth connection between the distributions in the midrange and the peak (tail) regions.[5] Compared with the groomed distributions with only the resummation on $\ln\rho$ (blue dot-dashed lines in Fig. 5), we observe that the resummation of the large logarithm $\ln y_c$ yields a significant enhancement through all the regions.

The red solid (dashed) lines in Fig. 5 are the plain jet mass distributions to an accuracy of NLL+NLO without (with) the NGL effects. Compared with the fully resummed groomed distributions, we see that the NGL effects are quite suppressed by the grooming process. Also, especially for the quark-initiated jet, we see that the jet mass distribution for $\rho \sim [10^{-3}, 0.1]$ is affected by the grooming. However, for the gluon-initiated jet, the grooming is not so effective. The gluon jet distributions are usually broad and have a relatively thick tail region. Hence, the grooming parameter $y_c = 0.1$ might not be large enough to suppress multiple soft or collinear-soft gluon emissions.

## VII. CONCLUSION

We have investigated the factorization of the groomed jet mass distribution over a wide range of the jet mass in the effective-theory approach. Distinct modes exist in the tail, midrange, and peak regions contributing to each factorized part. In the tail region, the collinear modes are enough to describe the groomed jet function, which coincides with the ungroomed jet function. In the midrange region with small jet mass, we also need the ultracollinear and csoft modes. In the peak region with very small jet mass, the ucsoft modes are also required. We apply the effective theories appropriate in these regions to obtain the factorized groomed jet mass distributions and resum the large logarithms on $\rho$ and/or $y_c$. By combining all the results, we are able to have a bird's-eye view of the groomed jet mass distribution over the whole range of the jet mass.

[4] Because the results at NLO in Eq. (31) involve only single logarithms, the resummed results at LL accuracy can be estimated as $\mathcal{O}(1)$ and are comparable to the results at NLL accuracy shown in Eqs. (48) and (58).

[5] The interpolation points are somewhat arbitrary. For example, if we try interpolation around $\rho \sim 0.06$ instead of $\rho \sim 0.08$, the shapes of the combined distributions vary slightly, giving some uncertainties which need to be fixed by experimental measurements.

The main issue is to implement the grooming procedure in theoretical calculations systematically. We have focused on how to treat the remaining particles, especially a single particle in a jet, when those particles which fail the grooming criterion are removed from the jet. The grooming, rather than the tagging, is chosen such that the remaining particles contribute to the groomed jet mass even when the grooming criterion is not satisfied. This prescribes the theoretical computation at NLO. When the criterion fails, the remaining single particle contributes to the $\delta(M_J^2)$ part, which includes the IR divergence in real emissions. However, it is cancelled by the virtual corrections, which also contribute to the $\delta(M_J^2)$ part. Therefore, the groomed jet mass distribution is IR safe and starts from $\alpha_s^0$. We focus on the theoretical issues here, and we will consider a detailed phenomenological analysis following the experimental setup closely, along with the inclusion of the nonperturbative effects, in a future work.

In Ref. 24, the factorization for groomed jet substructure was considered at NNLL accuracy. The focus was on the factorization of the hemisphere jet mass distributions, in which the detailed ingredients of the factorizations are different. However, the line of reasoning leading to the factorization was similar to the factorization in the peak region in our paper. In Refs. 25,26, a phenomenological analysis was performed on the groomed jet mass distributions in mMDT and soft drop in QCD with an estimate of the nonperturbative effects. The approach of Ref. 27 is closely related to our analysis of the peak region.

The important features of our paper compared to previous papers are the following: First, we elaborate on how to implement grooming theoretically in the effective-theory approach. Different modes are identified with different momentum scaling, and the overlap regions between different modes are disentangled using the zero-bin subtraction. The UV and the IR divergences are carefully treated to see if they are separated. After a nontrivial computation and a consistency check, we find that the grooming procedure yields IR-safe factorized parts while the tagging procedure does not at NLO. Using grooming, all the factorized parts are IR finite, and the RG equations can be applied to each of them to resum large logarithms. We here emphasize again that the IR safety or at least no mixture of IR and UV divergences (as in the case of the parton distribution functions) is the essential requirement to construct and solve the RG equations.

Second, in order to scan the entire possible range of the jet mass, we include the midrange region with $\rho \sim y_c \ll 1$. It is located between the peak and the tail regions. In this region, the ultracollinear, the csoft and the collinear modes give factorized contributions, which are different from those in the other regions. The new results for the factorized groomed jet mass distribution, which were obtained using the mMDT and the soft drop, are presented, and the resummation on $\ln y_c$, as well as $\ln\rho$, turns out to be the appreciable enhancements that persists in the peak region. In the numerical analysis, the effect of the resummation on $\ln\rho$ and $\ln y_c$ is appreciable though we put $y_c = 0.1$. This enhances the groomed jet mass distribution by about 50–80% compared to the result without resummation on $\ln y_c$. Also, the nonglobal logarithms are negligible, which is the characteristic of the mMDT and the soft drop.

The study of the jet substructure has become a mature subject along with the concurrent experimental analysis. The jet substructure can be investigated in various observables other than the jet mass distribution. These observables can be probed by using higher-order contributions, which include nonglobal and clustering logarithms, or by considering the behavior of the signal jets. The analytical comparison of the jet structure between the QCD jets and the signal jets will be a cornerstone to discover new physics through the study of jets. These features will be explored in a future work based on the approaches presented in this paper.

## APPENDIX A: LIST OF ALL THE FUNCTIONS IN THE GROOMED MASS DISTRIBUTION FUNCTIONS

Here, we list all the functions that appear in the text. Throughout this section, $y$ denotes the grooming parameter $y_c$. Also all the results in this section can be applied to the case of the soft drop with $\beta = 0$ when $y_c$ is replaced with $z_{\rm cut}/(1 - z_{\rm cut})$.

The functions $I_q(y)$ and $I_g(y)$ in Eq. (24) are given by

$$\begin{aligned} I_q(y) &= \ln^2 y + \frac{3}{2}\ln y + \frac{3(1-y)}{1+y} - 2\ln y \ln(1+y) \\ &\quad + 2\mathrm{Li}_2\Big(\frac{y}{1+y}\Big) - 2\mathrm{Li}_2\Big(\frac{1}{1+y}\Big), \\ I_g(y) &= \ln^2 y + \frac{11}{6}\ln y + \frac{1-y}{18(1+y)^3}(67+130y+67y^2) \\ &\quad - 2\ln y\ln(1+y) + 2\mathrm{Li}_2\Big(\frac{y}{1+y}\Big) - 2\mathrm{Li}_2\Big(\frac{1}{1+y}\Big) \\ &\quad + \frac{T_R n_f}{C_A}\Big[-\frac{1}{3}\ln y - \frac{1-y}{18(1+y)^3}(13+22y+13y^2)\Big]. \end{aligned} \tag{A1}$$

In the limit of small $y$, they become

$$\begin{aligned} I_q(y) &\to \ln^2 y + \frac{3}{2}\ln y + 3 - \frac{\pi^2}{3}, \\ I_g(y) &\to \ln^2 y + \frac{11}{6}\ln y + \frac{67}{18} - \frac{\pi^2}{3} \\ &\quad + \frac{T_R n_f}{C_A}\Big(-\frac{1}{3}\ln y - \frac{13}{18}\Big). \end{aligned} \tag{A2}$$

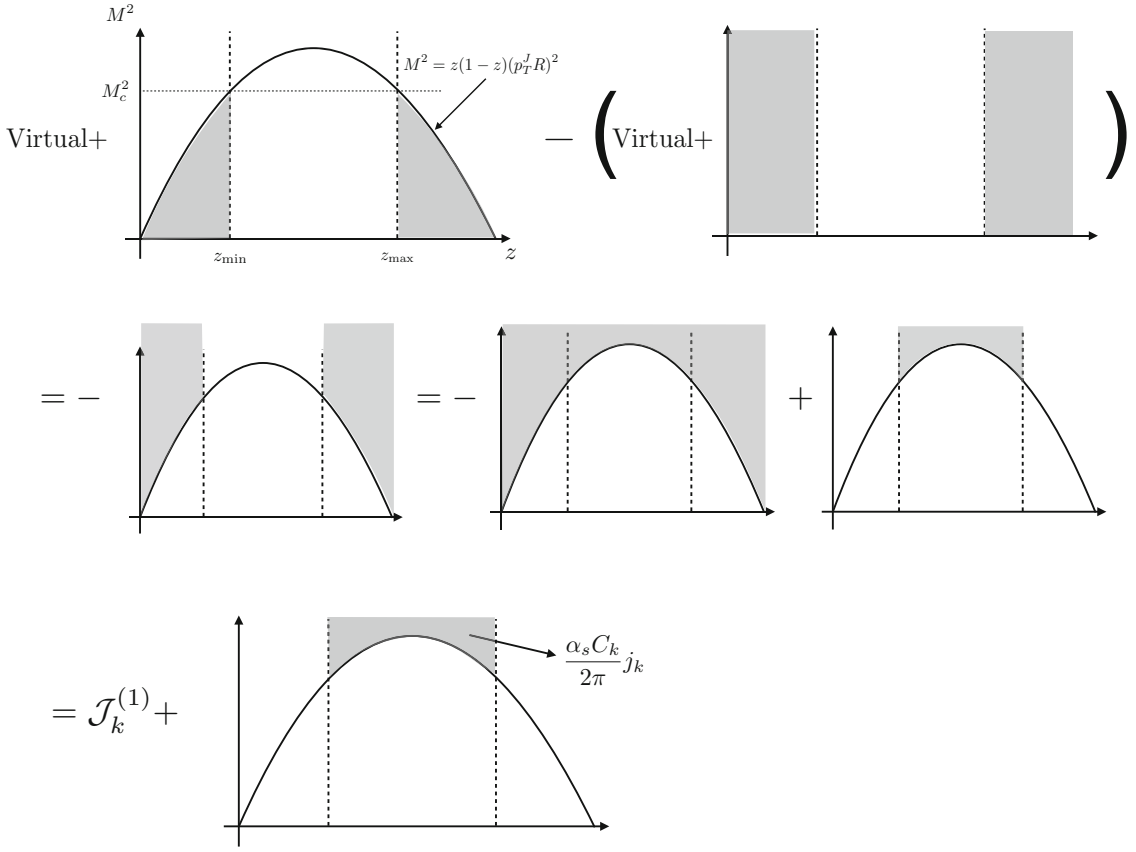


Fig. 6. Phase spaces for the collinear contribution in the region $\rho \ll y_c \sim \mathcal{O}(1)$ with the zero-bin subtraction. The resultant phase space is shown in the second row and it is decomposed into two parts. The first one yields the integrated jet function because it is the sum of the real contribution under the parabola with the virtual corrections. It is cancelled by performing the normalization, and the remaining second one yields the collinear function.

The functions $f_q(w)$ and $f_g(w)$ are given as

$$f_q(w) = -\frac{3}{2}w + 2\ln\frac{1+w}{1-w},$$
$$f_g(w) = -\frac{w}{12}(21+w^2) + 2\ln\frac{1+w}{1-w} + \frac{T_R n_f}{C_A}\frac{w}{12}(3+w^2). \tag{A3}$$

The functions $g_q(y)$ and $g_g(y)$ are given as

$$g_q(y) = \frac{3}{2} - \frac{3}{1+y} - 2\ln y,$$
$$g_g(y) = -\frac{1-y}{6(1+y)^3}(11+20y+11y^2) - 2\ln y\Big) + \frac{T_R n_f}{C_A}\frac{1-y^3}{3(1+y)^3}. \tag{A4}$$

In the limit $y \to 0$, they approach

$$g_q(y) \to -2\ln y - \frac{3}{2}, \quad g_g(y) \to -2\ln y - \frac{11}{6} + \frac{1}{3}\frac{T_R n_f}{C_A}. \tag{A5}$$

The functions $h_q(y)$ and $h_g(y)$ shown in Eqs. (22) and (30) are given as

$$\begin{aligned} h_q(y) = {} & -\frac{7}{2}\frac{1-y}{1+y} + \ln^2 y - 2\ln y\ln(1+y) - \frac{3y}{1+y}\ln y \\ & - \frac{3(1-y)}{1+y}\ln(1+y) + 2\mathrm{Li}_2\Big(\frac{1}{1+y}\Big) - 2\mathrm{Li}_2\Big(\frac{y}{1+y}\Big), \\ h_g(y) = {} & -\frac{1-y}{18(1+y)^3}(67+30y+67y^2) + \ln^2 y \\ & - 2\ln y\ln(1+y) - \frac{y}{3(1+y)^3}(12+21y+11y^2)\ln y \\ & - \frac{1-y}{3(1+y)^3}(11+20y+11y^2)\ln(1+y) \\ & + 2\mathrm{Li}_2\Big(\frac{1}{1+y}\Big) - 2\mathrm{Li}_2\Big(\frac{y}{1+y}\Big) \\ & + \frac{T_R n_f}{C_A}\Big[\frac{y}{3(1+y)^3}(3+3y+2y^2)\ln y \\ & + \frac{1-y^3}{9(1+y)^3}\Big(5+6\ln(1+y)\Big)\Big], \end{aligned} \tag{A6}$$

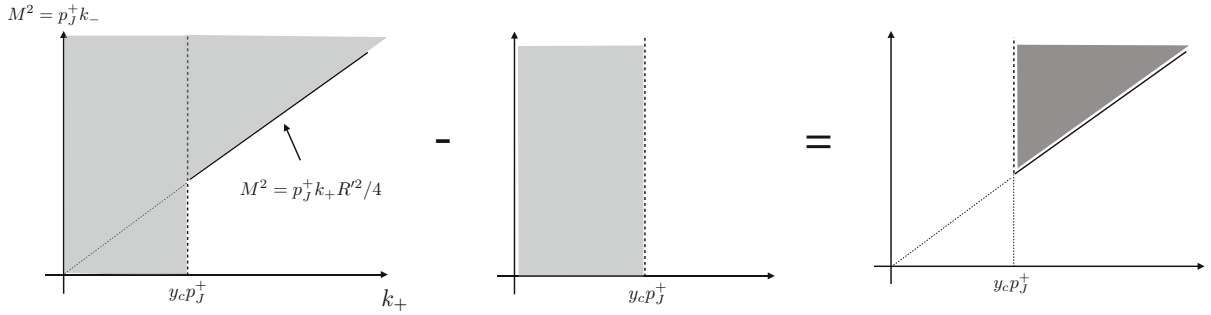


Fig. 7. Phase spaces for the csoft contributions with the zero-bin subtraction. The virtual contributions cancel because they are equal in the csoft and its zero-bin contributions. The csoft function can be obtained by integrating over the shaded region of the phase space.

and their limiting forms for small $y$ are given as

$$\begin{aligned} h_q(y) &\to -\frac{7}{2} + \frac{\pi^2}{3} + \ln^2 y, \\ h_g(y) &\to -\frac{67}{18} + \frac{\pi^2}{3} + \ln^2 y + \frac{5}{9}\frac{T_R n_f}{C_A}. \end{aligned} \tag{A7}$$

## APPENDIX B: CONVERSION OF THE Λ-DISTRIBUTION TO THE STANDARD PLUS DISTRIBUTION

The jet mass distributions and their factorized functions, such as the standard jet function and the csoft and the ucsoft functions with the Λ-distribution, can be expressed in terms of the standard plus functions. Let us define the dimensionless jet mass variable $\rho = M_J^2/Q^2$, where $Q = p_T^J R$. Then, the following Λ-distribution with a given function $g(M_J^2)$ can be written as

$$\begin{aligned} \Big[g(M_J^2)\Big]_{\Lambda^2} &= g(M_J^2) - \delta(M_J^2)\Big[\int_0^{\Lambda^2} \mathrm{d}M^2 g(M^2)\Big] \\ &= \frac{\hat{g}(\rho)}{Q^2} - \frac{\delta(\rho)}{Q^2}\Big[\int_0^{\Lambda^2/Q^2} \mathrm{d}\rho' g(\rho')\Big]. \end{aligned} \tag{B1}$$

Therefore, if we define the dimensionless function $\hat{g}(\rho) = Q^2 g(M_J^2)$, the distribution reads

$$\begin{aligned} Q^2\Big[g(M_J^2)\Big]_{\Lambda^2} &= \hat{g}(\rho) - \delta(\rho)\Big[\int_0^1 \mathrm{d}\rho' \hat{g}(\rho') - \int_{\Lambda^2/Q^2}^1 \mathrm{d}\rho' \hat{g}(\rho')\Big] \\ &= \Big[\hat{g}(\rho)\Big]_+ + \delta(\rho)\Big[\int_{\Lambda^2/Q^2}^1 \mathrm{d}\rho' \hat{g}(\rho')\Big]. \end{aligned} \tag{B2}$$

For example, the following Λ-distributions can be rewritten as

$$\begin{aligned} Q^2\Big[\frac{1}{M_J^2}\Big]_{\Lambda^2} &= \frac{1}{\rho_+} - \delta(\rho)\ln\frac{\Lambda^2}{Q^2}, \\ Q^2\Big[\frac{1}{M_J^2}\ln\frac{\mu^2}{M_J^2}\Big]_{\Lambda^2} &= \frac{1}{\rho_+}\ln\frac{\mu^2}{Q^2} - \Big(\frac{\ln\rho}{\rho}\Big)_+ \\ &\quad + \delta(\rho)\Big(\ln\frac{\Lambda^2}{Q^2}\ln\frac{\mu^2}{Q^2} - \frac{1}{2}\ln^2\frac{\Lambda^2}{Q^2}\Big). \end{aligned} \tag{B3}$$

## APPENDIX C: NLO CALCULATIONS OF $\tilde{\mathcal{C}}_K$ AND $\mathcal{S}_K^{II}$ WITH ZERO-BIN SUBTRACTION

The collinear contribution $\tilde{\mathcal{C}}_k(Q^2,\mu)$ in the limit $\rho \ll y_c \sim \mathcal{O}(1)$ can be computed by considering the phase spaces, as shown in Fig. 6. The collinear emission cannot satisfy the mMDT criterion; otherwise, $\rho$ becomes $\mathcal{O}(1)$. Therefore, it cannot contribute to the nonzero jet mass directly. In the first row in Fig. 6, the figure on the left-hand side shows the possible phase space for collinear emission along with virtual corrections. The figure on the right-hand side is the zero-bin subtraction, in which the ultracollinear modes contribute. Because the ultracollinear modes do not recognize the boundary of the ungroomed jet (the parabola), the phase space extends to infinity for $M^2$. Because the virtual corrections in the collinear and the zero-bin contributions are the same, they cancel through the subtraction and the available phase space is presented in the second row in Fig. 6.

The second figure in the second row yields the integrated jet function because it corresponds to the real emission under the parabola with the virtual correction. Finally, with the normalization dividing by $\mathcal{J}_k$, the collinear contribution with the zero-bin subtraction to NLO is given by

$$\frac{1}{\mathcal{J}_k}\Big(\mathcal{J}_k + \frac{\alpha_s C_k}{2\pi} j_k\Big) = 1 + \frac{\alpha_s C_k}{2\pi} j_k, \tag{C1}$$

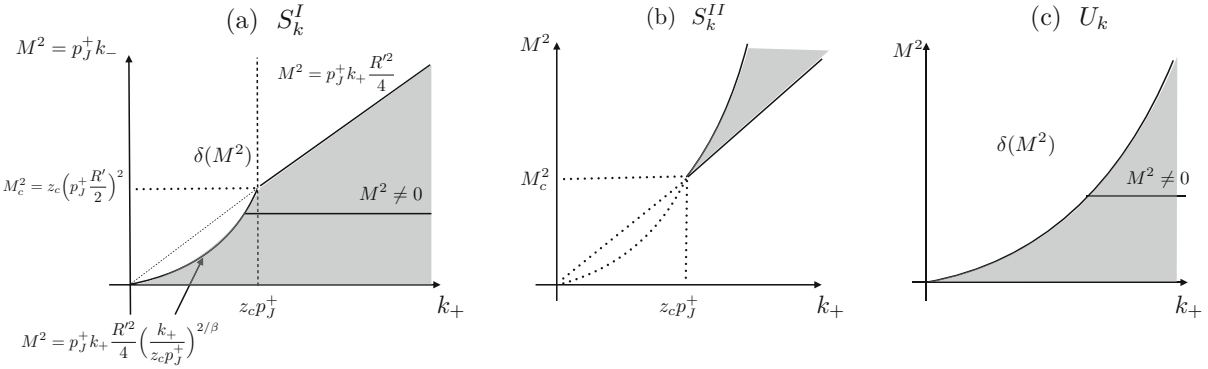


Fig. 8. Structures of the phase spaces: (a) $S_k^I(M^2)$ for the midrange region and (b) $\mathcal{S}_k^{II}$ and (c) $U_k(M^2)$ for the peak region. Here $z_c = z_{\rm cut}$. In (a) and (c), the phase spaces in the unshaded regions yield results proportional to $\delta(M^2)$ while only the shaded region in (b) contributes.

where $j_k$ is given by

$$j_k(y_c, M_c^2) = \Big(\frac{1}{\epsilon} + \ln\frac{\mu^2}{M_c^2}\Big) g_k(y_c) + h_k(y_c) - I_k(y_c), \quad \text{(C2)}$$

and the $g_k$, $h_k$ and $I_k$ are listed in Appendix A. From the structure of the phase space for $j_k$ in Fig. 6, the pole in Eq. (C2) is obviously an UV divergence. Therefore, the collinear contribution $\tilde{\mathcal{C}}_k(y_c, Q^2, \mu)$ is given by Eq. (30) after removing the UV divergence in Eq. (C2).

The csoft function $\mathcal{S}_k^{II}(y_c^2 Q^2, \mu)$ in the peak region ($\rho \ll y_c \ll 1$) can be computed by considering the phase spaces shown in Fig. 7. The first figure corresponds to the available phase space for the csoft gluon emission. Because the csoft radiation in this region cannot yield a nonzero groomed jet mass, it only contributes to the normalization of the jet mass distribution like $\tilde{\mathcal{C}}_k$. The second figure represents the phase space for the zero-bin subtraction. The virtual corrections in these two modes should be added, but they are the same. Therefore the virtual contributions cancel after the zero-bin subtraction. From the resultant phase space in Fig. 7, we easily see that no IR divergence and only the UV divergence exist.

The one loop result for the csoft contribution is obtained by integrating over the shaded region in the third figure in Fig. 7. The NLO results are given as

$$\mathcal{S}_k^{II}(y_c^2 Q^2, \mu) = 1 + \frac{\alpha_s C_k}{2\pi}\Big(\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\ln\frac{\mu^2}{y_c^2 Q^2} + \frac{1}{2}\ln^2\frac{\mu^2}{y_c^2 Q^2} - \frac{\pi^2}{12}\Big). \quad \text{(C3)}$$

After renormalization, we obtain the csoft function $\mathcal{S}_k^{II}(y_c^2 Q^2, \mu)$ in Eq. (39).

## APPENDIX D: NLO RESULTS FOR THE FACTORIZED FUNCTIONS FOR SOFT DROP WITH $\beta > 0$

We present the NLO results of the factorized functions for the groomed jet mass distributions using a soft drop with $\beta > 0$ in the limit $z_{\rm cut} \ll 1$. The soft drop condition is shown in Eq. (19). Because the same factorization formulae as in mMDT can be applied to a soft drop, Eq. (35) is the factorization theorem for the midrange region $\rho \sim z_{\rm cut} \ll 1$ and Eq. (38) is that for the peak region $\rho \ll z_{\rm cut} \ll 1$. Because the functions $\mathcal{C}_k$ and $J_k$ are not affected by the angular exponent $\beta$, we present $S_k^I$ in the midrange region and $\mathcal{S}_k^{II}$ and $U_k$ in the peak region. The result in the midrange region is new, and the results for the peak region have also been computed in Ref. 27, but we could not compare the results directly since the exact definition of the plus distribution function employed in Ref. 27 was not mentioned.

### 1. Midrange region: $\rho \sim z_c \ll 1$

The csoft function $S_k^I$ in the midrange region can be computed from the phase space illustrated in Fig. 8(a). For actual computation, we will consider the dimensionless csoft function $\bar{S}_k^I(\rho)$, where $\rho = M^2/Q^2$ and $Q = p_T^J R = p_J^+ R'/2$. The part with $\delta(\rho)$ comes from the virtual contribution, and the real contribution comes from the unshaded phase space in Fig. 8(a).

Also, the part with $\rho \neq 0$ can be obtained by integrating $k_+$ over the shaded region in Fig. 8(a). When $\rho > z_{\rm cut}$, the contribution is given as

$$M_{\rm cs}(\rho > z_{\rm cut}) = \frac{\alpha_s C_k}{\pi}\frac{1}{\rho}\Big(\frac{1}{\epsilon} + \ln\frac{\mu^2}{\rho^2 Q^2}\Big), \quad \text{(D1)}$$

and when $\rho \le z_{\rm cut}$, it is written as

$$M_{\rm cs}(\rho \le z_{\rm cut}) = \frac{\alpha_s C_k}{\pi} \frac{e^{\gamma_E \epsilon}}{\Gamma(1-\epsilon)} \frac{1}{\epsilon} \Big(\frac{\mu^2}{Q^2}\Big)^{\epsilon} z_{\rm cut}^{-\frac{2\epsilon}{2+\beta}} \rho^{-1-\epsilon\frac{2+2\beta}{2+\beta}}. \tag{D2}$$

Here, the poles in Eqs. (D1) and (D2) are the UV divergences. In Eq. (D2), an IR divergence occurs as $\rho \to 0$, which is associated with the part with $\delta(\rho)$ after employing the standard plus distribution.

Combining the results in Eqs. (D1) and (D2) with the parts proportional to $\delta(\rho)$ (*i.e.*, the virtual contribution and the real contribution from the unshaded region in Fig. 8(a)), we obtain the (bare) NLO result for $\bar{S}_k^I(\rho)$ as

$$\begin{aligned}\bar{S}_k^I(\rho,\mu) = \delta(\rho) + \frac{\alpha_s C_k}{2\pi}\Big\{ &\delta(\rho)\Big(-\frac{1}{\epsilon^2} - \frac{1}{\epsilon}\ln\frac{\mu^2}{Q^2} - \frac{1}{2}\ln^2\frac{\mu^2}{Q^2} \\ &+ \frac{2}{2+\beta}\ln^2 z_c + \frac{\pi^2}{12}\Big) \\ &+ \Big[\frac{2}{\rho}\Big(\frac{1}{\epsilon} + \ln\frac{\mu^2}{\rho^2 Q^2} + \frac{2}{2+\beta}\ln\frac{\rho}{z_c}\Big)\Big]_+ \\ &- \frac{2}{2+\beta}\frac{2}{\rho}\ln\frac{\rho}{z_{\rm cut}}\Theta(\rho - z_{\rm cut})\Big\}. \end{aligned} \tag{D3}$$

Here, the IR divergence as $\rho \to 0$ in Eq. (D2) is cancelled from the virtual correction, and all the remaining poles are UV divergences.

### 2. Peak region: $\rho \ll z_c \ll 1$

The csoft contribution in the peak region can be obtained from the phase space, as shown in Fig. 8(b). The phase space after the zero-bin subtraction is similar to the mMDT case (or the soft drop with $\beta = 0$), but the vertical line ($k_+ = y_c p_J^+$) in Fig. 7 is replaced by a curve with $\beta > 0$. From the phase space, the divergence obviously has an UV origin. Finally, the bare csoft function $\bar{\mathcal{S}}_k^{II}$ is given to NLO as

$$\begin{aligned}\bar{\mathcal{S}}_k^{II}(\mu) = 1 + \frac{\alpha_s C_k}{2\pi}\frac{1}{1+\beta}\Big(&\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\ln\frac{\mu^2}{z_{\rm cut}^2 Q^2} \\ &+ \frac{1}{2}\ln^2\frac{\mu^2}{z_{\rm cut}^2 Q^2} - \frac{\pi^2}{12}\Big). \end{aligned} \tag{D4}$$

For the ucsoft function, the phase space is shown in Fig. 8(c), in which the part with $\rho \neq 0$ comes from the shaded region while the rest of the phase space yields the $\delta(\rho)$ part. Following the method in computing $\bar{S}_k^I$, the dimensionless ucsoft functions $\bar{U}_k(\rho)$ ($= Q^2 U_k(M^2)$) are given to NLO as

$$\begin{aligned}\bar{U}_k(\rho,\mu) = \delta(\rho) + \frac{\alpha_s C_k}{2\pi}\Big\{ &-\frac{2+\beta}{1+\beta}\delta(\rho)\Big(\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\ln\frac{\mu^2}{Q^2 z_{\rm cut}^{\frac{2}{2+\beta}}} \\ &+ \frac{1}{2}\ln^2\frac{\mu^2}{Q^2 z_{\rm cut}^{\frac{2}{2+\beta}}} - \frac{\pi^2}{12}\Big) \\ &+ \Big[\frac{2}{\rho}\Big(\frac{1}{\epsilon} + \ln\frac{\mu^2}{Q^2 z_{\rm cut}^{\frac{2}{2+\beta}}} - \frac{2+2\beta}{2+\beta}\ln\rho\Big)\Big]_+\Big\}. \end{aligned} \tag{D5}$$

Here, all the poles are UV divergences. From Eqs. (D3), (D4) and (D5), we easily check that the relation $\bar{S}_k^I = \bar{S}_k^{II}\bar{U}_k$ holds to NLO in $\alpha_s$. Therefore, the results in the midrange and the peak regions are consistent.

## ACKNOWLEDGMENTS

J. Chay is supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant No. NRF-2016R1D1A1B03935799). C. Kim is supported by the Basic Science Research Program through the NRF funded by the Ministry of Science and ICT (Grant No. NRF-2017R1A2B4010511).

## REFERENCES

[1] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008), arXiv:0802.2470 [hep-ph].
[2] M. Dasgupta, A. Fregoso, S. Marzani and G. P. Salam, JHEP **09**, 029 (2013), arXiv:1307.0007 [hep-ph].
[3] A. J. Larkoski, S. Marzani, G. Soyez and J. Thaler, JHEP **05**, 146 (2014), arXiv:1402.2657 [hep-ph].
[4] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, Phys. Rev. D **79**, 074012 (2009), arXiv:0810.0934 [hep-ph].
[5] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa and S. Marzani, JHEP **08**, 064 (2010), arXiv:1004.3483 [hep-ph].
[6] H-n. Li, Z. Li and C. P. Yuan, Phys. Rev. Lett. **107**, 152001 (2011), arXiv:1107.4535 [hep-ph].
[7] H-n. Li, Z. Li and C. P. Yuan, Phys. Rev. D **87**, 074025 (2013), arXiv:1206.1344 [hep-ph].
[8] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani and M. Spannowsky, JHEP **10**, 126 (2012), arXiv:1207.1640 [hep-ph].
[9] Y-T. Chien, R. Kelley, M. D. Schwartz and H. X. Zhu, Phys. Rev. D **87**, 014010 (2013), arXiv:1208.0010 [hep-ph].
[10] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, Phys. Rev. D **88**, 054031 (2013), arXiv:1302.0846 [hep-ph].
[11] Z. L. Liu, C. S. Li, J. Wang and Y. Wang, JHEP **04**, 005 (2015), arXiv:1412.1337 [hep-ph].
[12] C. W. Bauer, S. Fleming and M. E. Luke, Phys. Rev. D **63**, 014006 (2000), arXiv:hep-ph/0005275 [hep-ph].

[13] C. W. Bauer, S. Fleming, D. Pirjol and I. W. Stewart, Phys. Rev. **D63**, 114020 (2001), arXiv:hep-ph/0011336 [hep-ph].
[14] C. W. Bauer, D. Pirjol and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002), arXiv:hep-ph/0109045 [hep-ph].
[15] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein and I. W. Stewart, Phys. Rev. **D66**, 014017 (2002), arXiv:hep-ph/0202088 [hep-ph].
[16] Y-T. Chien, A. Hornig and C. Lee, Phys. Rev. D **93**, 014033 (2016), arXiv:1509.04287 [hep-ph].
[17] A. Hornig, Y. Makris and T. Mehen, JHEP **04**, 097 (2016), arXiv:1601.01319 [hep-ph].
[18] D. W. Kolodrubetz, P. Pietrulewicz, I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, JHEP **12**, 054 (2016), arXiv:1605.08038 [hep-ph].
[19] A. Idilbi and C. Kim, J. Korean Phys. Soc. **73**, 1230 (2018), arXiv:1606.05429 [hep-ph].
[20] Z-B. Kang, K. Lee and F. Ringer, JHEP **04**, 110 (2018), arXiv:1801.00790 [hep-ph].
[21] M. Aaboud *et al.* (ATLAS), Phys. Rev. Lett. **121**, 092001 (2018), arXiv:1711.08341 [hep-ex].
[22] M. Dasgupta, A. Fregoso, S. Marzani and A. Powling, Eur. Phys. J. **C73**, 2623 (2013), arXiv:1307.0013 [hep-ph].
[23] C. Frye, A. J. Larkoski, M. D. Schwartz and K. Yan, (2016), arXiv:1603.06375 [hep-ph].
[24] C. Frye, A. J. Larkoski, M. D. Schwartz and K. Yan, JHEP **07**, 064 (2016), arXiv:1603.09338 [hep-ph].
[25] S. Marzani, L. Schunk and G. Soyez, JHEP **07**, 132 (2017), arXiv:1704.02210 [hep-ph].
[26] S. Marzani, L. Schunk and G. Soyez, Eur. Phys. J. **C78**, 96 (2018), arXiv:1712.05105 [hep-ph].
[27] Z-B. Kang, K. Lee, X. Liu and F. Ringer, JHEP **10**, 137 (2018), arXiv:1803.03645 [hep-ph].
[28] B. Jager, M. Stratmann and W. Vogelsang, Phys. Rev. D **70**, 034010 (2004), arXiv:hep-ph/0404057 [hep-ph].
[29] M. Procura and I. W. Stewart, Phys. Rev. **D81**, 074009 (2010), [Erratum: Phys. Rev.D83,039902(2011)], arXiv:0911.4980 [hep-ph].
[30] A. Jain, M. Procura and W. J. Waalewijn, JHEP **05**, 035 (2011), arXiv:1101.4953 [hep-ph].
[31] M. Ritzmann and W. J. Waalewijn, Phys. Rev. D **90**, 054029 (2014), arXiv:1407.3272 [hep-ph].
[32] M. Dasgupta, F. Dreyer, G. P. Salam and G. Soyez, JHEP **04**, 039 (2015), arXiv:1411.5182 [hep-ph].
[33] M. Dasgupta, F. A. Dreyer, G. P. Salam and G. Soyez, JHEP **06**, 057 (2016), arXiv:1602.01110 [hep-ph].
[34] T. Kaufmann, A. Mukherjee and W. Vogelsang, Phys. Rev. D **92**, 054015 (2015), arXiv:1506.01415 [hep-ph].
[35] Z-B. Kang, F. Ringer and I. Vitev, JHEP **10**, 125 (2016), arXiv:1606.06732 [hep-ph].
[36] L. Dai, C. Kim and A. K. Leibovich, Phys. Rev. D **94**, 114023 (2016), arXiv:1606.07411 [hep-ph].
[37] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
[38] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993), arXiv:hep-ph/9305266 [hep-ph].
[39] M. Cacciari, G. P. Salam and G. Soyez, JHEP **04**, 063 (2008), arXiv:0802.1189 [hep-ph].
[40] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, JHEP **08**, 001 (1997), arXiv:hep-ph/9707323 [hep-ph].
[41] M. Procura and W. J. Waalewijn, Phys. Rev. D **85**, 114041 (2012), arXiv:1111.6605 [hep-ph].
[42] Z-B. Kang, F. Ringer and I. Vitev, JHEP **11**, 155 (2016), arXiv:1606.07063 [hep-ph].
[43] W. M-Y. Cheung, M. Luke and S. Zuberi, Phys. Rev. D **80**, 114021 (2009), arXiv:0910.2479 [hep-ph].
[44] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig and C. Lee, JHEP **11**, 101 (2010), arXiv:1001.0014 [hep-ph].
[45] J. Chay, C. Kim and I. Kim, Phys. Rev. **D92**, 034012 (2015), arXiv:1505.00121 [hep-ph].
[46] C. W. Bauer, F. J. Tackmann, J. R. Walsh and S. Zuberi, Phys. Rev. D **85**, 074006 (2012), arXiv:1106.6047 [hep-ph].
[47] M. Procura, W. J. Waalewijn and L. Zeune, JHEP **02**, 117 (2015), arXiv:1410.6483 [hep-ph].
[48] T. Becher, M. Neubert, L. Rothen and D. Y. Shao, Phys. Rev. Lett. **116**, 192001 (2016), arXiv:1508.06645 [hep-ph].
[49] L. Dai, C. Kim and A. K. Leibovich, Phys. Rev. D **95**, 074003 (2017), arXiv:1701.05660 [hep-ph].
[50] X. Liu, S-O. Moch and F. Ringer, Phys. Rev. Lett. **119**, 212001 (2017), arXiv:1708.04641 [hep-ph].
[51] G. P. Korchemsky and G. F. Sterman, Phys. Lett. **B340**, 96 (1994), arXiv:hep-ph/9407344 [hep-ph].
[52] R. Akhoury and I. Z. Rothstein, Phys. Rev. D **54**, 2349 (1996), arXiv:hep-ph/9512303 [hep-ph].
[53] S. W. Bosch, B. O. Lange, M. Neubert and G. Paz, Nucl. Phys. **B699**, 335 (2004), arXiv:hep-ph/0402094 [hep-ph].
[54] M. Dasgupta and G. P. Salam, Phys. Lett. **B512**, 323 (2001), arXiv:hep-ph/0104277 [hep-ph].
[55] A. Banfi, G. Marchesini and G. Smye, JHEP **08**, 006 (2002), arXiv:hep-ph/0206076 [hep-ph].
[56] A. V. Manohar and I. W. Stewart, Phys. Rev. D **76**, 074002 (2007), arXiv:hep-ph/0605001 [hep-ph].
[57] G. P. Korchemsky and A. V. Radyushkin, Nucl. Phys. **B283**, 342 (1987).
[58] I. A. Korchemskaya and G. P. Korchemsky, Phys. Lett. **B287**, 169 (1992).
[59] M. Neubert, Phys. Rev. D **72**, 074025 (2005), arXiv:hep-ph/0506245 [hep-ph].
[60] T. Becher and M. Neubert, Phys. Rev. Lett. **97**, 082001 (2006), arXiv:hep-ph/0605050 [hep-ph].