


Dynamical grooming of QCD jets

Yacine Mehtar-Tani^{1,*}, Alba Soto-Ontoso^{1,†} and Konrad Tywoniuk^{2,‡}

¹*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*

²*Department of Physics and Technology, University of Bergen, 5007 Bergen, Norway*

 (Received 12 November 2019; accepted 21 January 2020; published 6 February 2020)

We propose a new class of infrared-collinear and Sudakov safe observables with an associated jet grooming technique that removes dynamically soft and large angle branches. It is based on identifying the *hardest* branch in the Cambridge/Aachen reclustering sequence and discarding prior splittings that occur at larger angles. This leads to a dynamically generated cutoff on the phase space of the tagged splitting that is encoded in a Sudakov form factor. In this exploratory study we focus on the mass and momentum sharing distributions of the tagged splitting which we analyze analytically to modified leading-logarithmic accuracy and compare to Monte Carlo simulations.

DOI: [10.1103/PhysRevD.101.034004](https://doi.org/10.1103/PhysRevD.101.034004)

I. INTRODUCTION

Jets, or collimated sprays of particles originating from the fragmentation of energetic quarks and gluons, are among the most prominent features of high-energy particle collisions. The analysis of jet observables is crucial to study the theory of strong interactions, QCD, in the perturbative regime, including the running of the strong coupling constant α_s . These phenomena also play an important role in constraining the background in searches for heavy particles, including the Higgs boson [1] and particles beyond the Standard Model [2].

In the case of high-energy hadronic collisions, however, the observables are strongly affected by a wide range of processes that are hard to account for in perturbation theory and conventional resummation techniques. These include radiation from outside of the jet (nonglobal logarithms) [3] and nonperturbative effects such as hadronization and underlying event activity. In the last decade, tackling these challenges has led to an improved analytical understanding of jet substructure (see [4–6] for recent reviews), coinciding with the maturing of fast and versatile jet reclustering procedures [7–9].

In this context, several jet grooming techniques designed to reduce the jet’s sensitivity to nonlocal and nonperturbative physics have been developed. Such techniques have

further evolved toward being easier to interpret in terms of perturbative QCD [10,11]. Representative examples are the modified MassDrop tagger (mMDT) grooming [10] and SoftDrop (SD) grooming [12] that provide a two-parameter algorithm to determine the first branching in an angular-ordered tree that is deemed to be sufficiently perturbative. Given the i th primary emission off an angular-ordered jet (corresponding to the i th branch of a Cambridge/Aachen reclustering along the leading flow of energy), where $p_{T,i_1} > p_{T,i_2}$ are the energies of the two splitting products, one removes such emissions until one identifies the first whose momentum sharing fraction satisfies the condition $z > z_{\text{cut}}(\theta/R)^\beta$. Grooming recursively along the primary and secondary emission branches strongly reduces non-perturbative effects in specific cases [13]. Other techniques, such as trimming [14], recluster the jet with a smaller cone size and remove substructures below a certain energy cutoff. Grooming techniques have also proven promising to study the internal structure of quenched jets in the context of heavy-ion collisions [15–18]; see also [19].

Furthermore, substructure techniques probe our knowledge of the multiparticle regime of QCD. Thus, designing new observables often goes hand in hand with a grooming scheme that permits a direct comparison to experimental data. For example, the momentum sharing variable z_g of the first accepted emission in SoftDrop with $\beta = 0$, that coincides with mMDT grooming [10], turns out to be an ultraviolet fixed point [20] so that its distribution does not depend on the strong coupling constant. Other examples, such as the groomed jet mass [21,22], have been calculated to next-to-leading-logarithmic accuracy.

Despite the many successes, current jet substructure techniques are often quite simple but lacking an internal “logic” that would allow us to estimate the most natural choice for the grooming parameters. These procedures are

*mehtartani@bnl.gov

†ontoso@bnl.gov

‡konrad.tywoniuk@uib.no

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP³.

sensitive to the choice of parameters, e.g., z_{cut} and β in the case of SD, and their optimal values, in terms of resilience to the underlying event or other distortions, can possibly depend on jet p_T , underlying event activity and other unknown parameters. Clearly, if $z_{\text{cut}} \ll 1$ the sensitivity to nonperturbative infrared effects is enhanced. Moreover, from an analytical point of view, their inclusion generates new scales on the level of jet substructure observables that complicate the understanding of the different contributing modes. This appears because the intrinsic jet scale is fluctuating on a jet-by-jet basis. One would therefore wish for a method that aligns more closely with the intrinsic properties of a given jet without the need for fine tuning.

In this work, we aim to alleviate some of these shortcomings. We consider a class of observables based on selecting, or tagging, the *hardest* splitting in an angular-ordered shower, where the “hardness” is characterized by a pseudo-energy-correlation variable as follows. Given the i th $1 \rightarrow 2$ splitting in the Cambridge/Aachen reclustering sequence, the variable that measures the hardness of a jet, or in other words defines the “hardest” splitting within a jet, is defined as

$$\kappa^{(a)} = \frac{1}{p_T} \max_{i \in \text{C/A seq.}} \left[z_i (1 - z_i) p_{T,i} \left(\frac{\theta_i}{R} \right)^a \right], \quad (1)$$

where p_T and R are the energy and radius of the jet respectively, z_i is the momentum sharing fraction, $p_{T,i}$ the energy of the parent, θ_i the relative angle of the splitting and a is a free parameter whose physical interpretation will be discussed below.

Note that for $a = 2$, we would select the splitting with the shortest formation time $t_f^{-1} \sim \kappa^{(2)} p_T$. We refer to this case as TimeDrop in what follows. Alternatively, for $a = 1$ we tag the branching with the largest relative transverse momentum $k_T \sim \kappa^{(1)} p_T$ and name this option k_T Drop. Finally, $a = 0$ corresponds to the splitting with the most symmetric momentum sharing and is called z Drop in what follows. In fact, $a = 0$ leads to collinear sensitivity (see the Appendix A), and we will rather use $a = 0.1$ for all practical purposes below.

Having identified a genuinely hard branching in the shower, we suggest two strategies.

- (i) In tagging mode, the kinematics of the hardest splitting informs the observable one wishes to compute. This will be the main focus of this paper.
- (ii) In grooming mode, one discards all emissions taking place prior to the hard splitting in the reclustering sequence. This procedure can easily be iterated along all the branches of the jet. We will pursue this strategy for other groomed observables in an upcoming publication.

The main advantage of this method is that it autogenerates the conditions for tagging or grooming on a jet-by-jet basis. While a similar strategy is also pursued within e.g.,

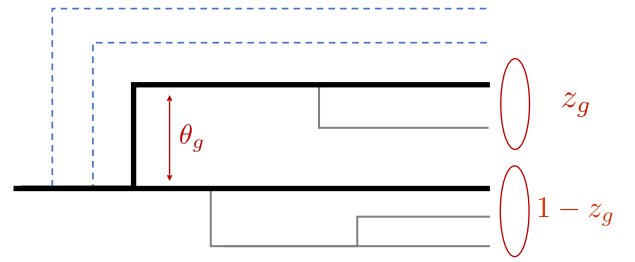


FIG. 1. Dynamical grooming applied to an angular-ordered tree. The splitting represented by the thick (black) line has the largest $\kappa^{(a)}$ in the tree. In tagging mode, observables are calculated using the kinematic variables z_g and θ_g of the tagged splitting. In grooming mode, softer splittings that appear earlier in the tree, i.e., at larger angles, are discarded and the jet kinematics is adjusted accordingly.

jet pruning [23], the procedure in our case is simpler to implement and closer in spirit to the physics of color coherence. In fact, softer emissions in the C/A sequence prior to the hardest one can be considered as radiation off the total charge. Our procedure only depends on one parameter which defines what we mean by the hardest emission inside a jet in contrast to most other techniques that involve two (extrinsic) parameters. Hence, we refer to this procedure as dynamical grooming. A schematic illustration of how to dynamically groom an angular-ordered shower can be found in Fig. 1.

This paper is structured as follows. In Sec. II, we discuss vetoed showers and introduce the probability for a splitting to be the hardest. In Sec. III, we employ the derived Sudakov form factor to compute a family of observables based on the tagged splitting. We then perform analytical calculations in the modified leading-logarithmic approximation (MLLA) for the tagged mass and z distributions in Secs. III A and III B. Finally, Sec. IV is dedicated to Monte Carlo simulations of proton-proton collisions. We generate and analyze the Lund planes for the tagged splittings within the different dynamical grooming settings and present a systematic study of the impact of nonperturbative phenomena on the tagged mass and z distributions. We end with a short discussion and outlook in Sec. V.

II. VETOED SHOWERS AND TAGGING

Dynamically grooming a jet amounts to a certain reorganization of the conventional parton shower, where we will assume angular ordering. Given a specific $1 \rightarrow 2$ splitting in the jet history, the procedure forces the emissions taking place both before, i.e., at angles larger than the selected splitting, and after, i.e., at smaller angles, not to allow for a harder emission. The reorganization depends on the properties of a given jet allowing for a procedure that is more adapted to account for jet-by-jet fluctuations.

The hardness variable $\kappa^{(a)}$ is easily accessible in experimental data together with Monte Carlo showers, and will

be used in Sec. IV. In order to tackle the problem analytically, in a transparent manner and up to the required logarithmic precision, some simplifications are in order. To next-to-leading-logarithmic accuracy in an angular-ordered shower, it can be shown that the hardest splitting takes place off the leading particle in the jet or, in other words, on the primary Lund plane [24,25]. In this case, we neglect the energy depletion of the leading jet $p_{T,i} \simeq p_T$ and hence explicit dependence on momenta cancel out in Eq. (1). Accordingly, for a collinear safe definition, i.e., for $a > 0$ in Eq. (1), the hardness of a tagged splitting is given by

$$\kappa^{(a)} = z(1-z) \left(\frac{\theta}{R} \right)^a. \quad (2)$$

Whenever it is obvious from the context, we will simply write $\kappa \equiv \kappa^{(a)}$.

The central quantity in our computations is the following Sudakov form factor:

$$\ln \Delta(\kappa|a) = - \int_0^R \frac{d\theta}{\theta} \int_0^1 dz \frac{\alpha_s(k_t^2)}{\pi} P(z) \times \Theta(z(1-z)(\theta/R)^a > \kappa), \quad (3)$$

where $P(z)$ is the splitting function and the Heaviside function vetoes emissions with a $\kappa^{(a)}$ larger than the measured, or tagged, emission. Note that the angular integral, spanning between 0 and R , enforces this veto all over the primary Lund plane of the jet. Finally, $\alpha_s(k_t^2)$ is the strong coupling constant evaluated at $k_t = z(1-z)\theta p_T$, the transverse momentum generated at the splitting.

Such a form factor arises as a remainder contribution from the vetoed showers occurring before and after the hard emission. A similar construction was previously used as a method to match parton showers with next-to-leading order contributions where the hardest emission was the one with the largest $k_t \sim \kappa^{(1)}$ [26]. Here, we will proceed with a more direct line of reasoning to derive the relevant probability distribution. Taking the derivative and then integrating over κ in Eq. (3) leads to the following identity:

$$\int_0^\infty d\kappa \frac{d}{d\kappa} \Delta(\kappa) = \Delta(\infty) - \Delta(0) \quad (4)$$

where $\Delta(\kappa) \equiv \Delta(\kappa|a)$. Clearly, we have that $\Delta(\infty) = 1$. For $a > 0$, the Sudakov given in Eq. (3) is infrared and collinear finite. Therefore, we can safely take the limit $\kappa \rightarrow 0$, resulting in $\Delta(0) = 0$. This is no longer the case for collinear unsafe observables, explicitly for $a = 0$, where we have to introduce a nonperturbative cutoff scale to regulate the integrals. We will treat this particular case in more detail in Appendix A, and focus in the remainder of the paper on the collinear safe taggers.

Then, for collinear safe observables we can construct a normalized probability distribution of the splitting with the

largest $\kappa^{(a)}$ in an angular-ordered shower. Owing to the fact that

$$\frac{d}{d\kappa} \Delta(\kappa) = \int_0^R \frac{d\theta}{\theta} \int_0^1 dz \frac{\alpha_s(k_t^2)}{\pi} P(z) \Delta(\kappa^{(a)}|a) \times \delta(z(1-z)(\theta/R)^a - \kappa), \quad (5)$$

it follows that

$$\int_0^R \frac{d\theta}{\theta} \int_0^1 \frac{dz}{z} P(z, \theta) = 1, \quad (6)$$

where

$$\mathcal{P}(z, \theta) = \frac{\alpha_s(k_t^2)}{\pi} z P(z) \Delta(\kappa|a) \quad (7)$$

is the probability of splitting giving rise to the momentum sharing fraction z at angle θ that results in the largest $\kappa^{(a)}$ in the shower, i.e., to be the hardest splitting.

In order to gain analytic insight, we will work for simplicity in the modified leading-logarithmic approximation that assumes angular ordering and where the one loop Altarelli-Parisi quark-gluon splitting function can be approximated as

$$P_{gq}(z) = C_F \frac{1 + (1-z)^2}{z} \approx \frac{2C_F}{z} \left(1 - \frac{3}{4}z \right), \quad (8)$$

where the second term comes from approximating the finite part of the splitting function by $\int_0^1 dz [P(z)/C_F - 2/z] = -3/2$. This corresponds to the modified leading-logarithmic approximation. For fixed coupling and using the approximation in Eq. (8), the integrals in Eq. (3) can be done analytically. The final expression for the Sudakov form factor in the MLLA is then

$$\ln \Delta(\kappa|a) = -\frac{\bar{\alpha}}{a} \left[\ln^2 \kappa + \frac{3}{2} (\ln \kappa + 1 - \kappa) \right], \quad (9)$$

where $\bar{\alpha} \equiv \alpha_s C_F / \pi$. Only the first two terms are relevant to leading-logarithmic (LL) accuracy. For the numerical evaluations, we will keep the term of order 1 to ensure exact normalization of our observables. An analogous derivation of the Sudakov form factor for gluon-initiated jets can be found in Appendix C.

III. COMPUTING TAGGED OBSERVABLES

Since our procedure exploits the properties of the hardest branching in the jet shower, we will be interested in observables that are of the same form as Eq. (2). Generically, such distributions are then given by

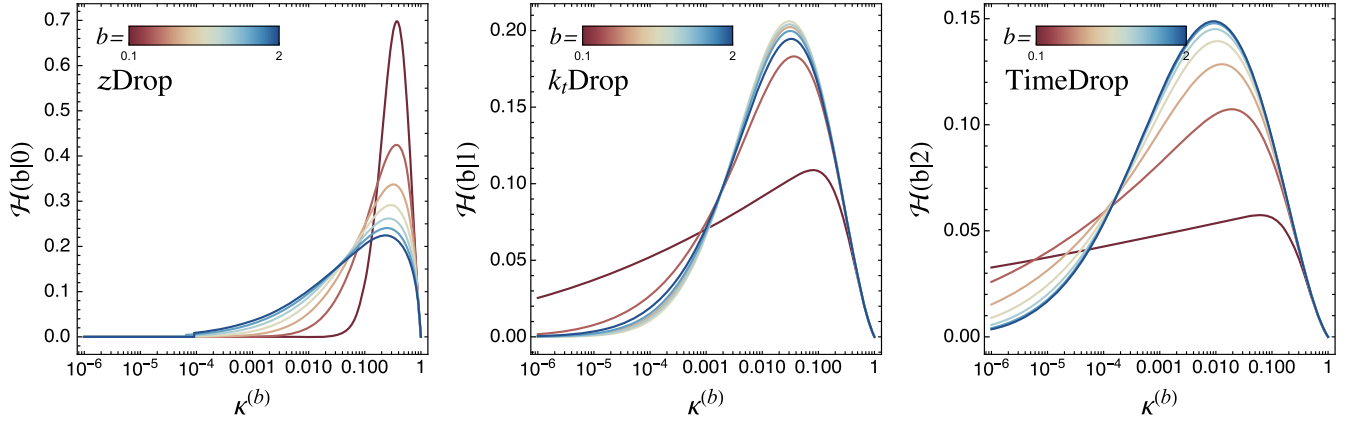


FIG. 2. The normalized distribution $\mathcal{H}(b|a)$ for three grooming modes: $z\text{Drop}$ ($a = 0.1$, left panel); $k_t\text{Drop}$ ($a = 1$, central panel); and TimeDrop ($a = 2$, right panel). We have generated distributions for a range of b values in Eq. (12) for fixed coupling at MLLA.

$$\frac{1}{\sigma} \frac{d\sigma}{d\kappa^{(b)}} \Big|_a = \int_0^R \frac{d\theta}{\theta} \int_0^1 dz \frac{\alpha_s(k_t^2)}{\pi} P(z) \Delta(\kappa^{(a)}|a) \times \delta(z(1-z)(\theta/R)^b - \kappa^{(b)}), \quad (10)$$

where $\sigma|_a$ represents the fact that the cross section distribution is measured given a tagged splitting with the largest $\kappa^{(a)}$ in the parton shower. The normalized distribution of the jet hardness is then written as

$$\mathcal{H}(b|a) \equiv \frac{1}{\sigma} \frac{d\sigma}{d\kappa^{(b)}} \Big|_a. \quad (11)$$

Hence, the first argument of the function is related to the observable $\kappa^{(b)}$ that is measured on the kinematics of the tagged splitting while the second defines what we mean by tagging the hardest splitting; i.e., we identify the splitting with the largest $\kappa^{(a)}$. Whenever it is obvious from the context, we will simply denote $\kappa^{(b)} \equiv \kappa$.

Without losing generality, we work in the MLL approximation and treat $z \ll 1$. Then, assuming $b > 0$, we evaluate the δ -function to obtain

$$\mathcal{H}(b|a) = \frac{1}{b} \int_{\kappa}^1 dz \frac{\alpha_s(k_t^2)}{\pi} P(z) \Delta(z^{1-\frac{a}{b}} \kappa^{\frac{a}{b}}|a), \quad (12)$$

where now the argument of the running coupling is $k_t = z^{1-\frac{1}{b}} \kappa^{\frac{1}{b}} Q$ with $Q \equiv p_T R$ being the jet scale. It is straightforward to check that Eq. (12) is normalized to unity. After further simplifications, assuming $a \neq b$, the final expression reads

$$\mathcal{H}(b|a) = \frac{1}{b-a} \int_{\kappa}^1 \frac{dx}{x} \frac{\alpha_s(k_t^2)}{\pi} \tilde{P}(x^{\frac{b}{b-a}} \kappa^{-\frac{a}{b-a}}) \Delta(x|a), \quad (13)$$

where we introduced the notation $\tilde{P}(z) \equiv zP(z)$ and where now $k_t = x^{\frac{b-1}{b-a}} \kappa^{\frac{1-a}{b-a}} Q$.

It is clear from Eq. (13) that, except when $b = 0$, $\mathcal{H}(b|a)$ is an infrared and collinear safe quantity since it admits a Taylor expansion in the coupling constant. For $b = 0$, the x integration goes from 0 to κ and the integrand exhibits an essential singularity at 0 that is regulated by the Sudakov form factor. The resulting integral is an asymptotic series, and hence finite, although the perturbative expansion is divergent term by term. In this case, $\mathcal{H}(0|a)$ is said to be Sudakov safe [20,27].

Remarkably, in the double logarithmic approximation (DLA), where we drop the second term in Eq. (8), and with fixed coupling, we obtain

$$\begin{aligned} \mathcal{H}(b|a) &\simeq \frac{2\bar{\alpha}}{b-a} \int_{\kappa}^1 \frac{dx}{x} e^{-\frac{a}{b} \ln^2 x} \\ &\simeq \frac{\sqrt{\pi\bar{\alpha}a}}{b-a} \left[\text{erf}\left(\sqrt{\frac{\bar{\alpha}a}{b^2}} \ln \kappa\right) - \text{erf}\left(\sqrt{\frac{\bar{\alpha}}{a}} \ln \kappa\right) \right], \end{aligned} \quad (14)$$

which turns out to be invariant under the transformation $a \rightarrow b^2/a$, i.e.,

$$\mathcal{H}(b|a) \approx \mathcal{H}\left(b \left| \frac{b^2}{a} \right.\right), \quad (15)$$

where the equality only holds at DLA. This immediately singles out $a = b$ as a ‘‘fixed point’’ of this class of observables. The dual distributions correspond to different tagging modes and in the limiting case $a \rightarrow 0$ this amounts to strong and weak grooming in the left- and right-hand sides of Eq. (15), respectively. Below we will come back to the meaning of these observations.

We plot the normalized distribution $\mathcal{H}(b|a)$ as function of the variable $\kappa^{(b)}$ in Fig. 2 for three values of a , corresponding (from left to right) to $z\text{Drop}$ ($a \approx 0$), $k_t\text{Drop}$ ($a = 1$) and TimeDrop ($a = 2$). For each grooming setting we plot the distributions for different variables $\kappa^{(b)}$ for $0.1 < b < 2$. To observe the approximate duality Eq. (15) it

is sufficient to notice that in the central panel (k_t Drop) the distributions reach the maximal values at $b = a$ and start decreasing at $b > a$. For completeness, similar results for gluon-initiated jets are displayed in Appendix C.

In order to discuss the qualitative features of the spectrum at this level it is sufficient to focus on the case $b > a > 0$. The asymptotic behavior of the κ -distribution, i.e., $\ln^2 \kappa \gg (\bar{a}a)^{-1}$, reads

$$\mathcal{H}(b|a) \approx \frac{b}{b-a} \frac{1}{\ln \kappa^{-1}} e^{-\frac{\bar{a}a}{b^2} \ln^2 \kappa}. \quad (16)$$

We observe two qualitative features. First, the distribution peaks at $\ln \kappa \sim (\bar{a}a/b^2)^{-1/2}$. Hence, the peak of the distribution shifts to larger values of κ with decreasing a . Second, from Eq. (16), we see that the distribution flattens as a decreases. The opposite case, i.e., $a > b > 0$, can be found in a similar way, e.g., using Eq. (15) at DLA. These features are seen in Fig. 2. As a result, the limits $a \rightarrow \infty$ and $a \rightarrow 0$ exhibit similar behavior despite their different groomed modes.

Turning now to the special case, when $a = b > 0$, we evaluate Eq. (12) which gives

$$\mathcal{H}(a|a) = \frac{1}{a} \Delta(\kappa|a) \int_{\kappa}^1 dz \frac{\alpha_s(k_t^2)}{\pi} P(z), \quad (17)$$

where now $k_t = z^{1-\frac{1}{a}} \kappa^{\frac{1}{a}} Q$. The remaining integral in Eq. (17) is regulated from below and is finite. In the MLLA and at fixed coupling we get

$$\mathcal{H}(a|a) = -\frac{2\bar{\alpha}}{a} \left[\ln \kappa + \frac{3}{4}(1-\kappa) \right] \Delta(\kappa|a). \quad (18)$$

It turns out that the distribution measured this way corresponds to the plain distribution to LL accuracy, i.e., the distribution of the observables without any grooming; see Eq. (20) for a concrete example. Hence, having $b \sim a$ corresponds to a low degree of grooming, such that the distribution closely resembles the plain one, while $b \gg a$ results in strong grooming.

We now proceed to consider in more detail two important observables in jet physics, namely the mass and momentum sharing fraction z properties of the tagged, hard splitting. As a further example, we will consider the tagged k_t distribution in Appendix B.

A. Tagged mass distribution ($b=2$)

The case $b=2$ is related to the mass of a given splitting or, in other words, to the virtuality m^2 of the parent particle that decayed. Defining the rescaled variable $\rho \equiv m^2/(p_T R)^2$, the normalized distribution is simply $\mathcal{H}(2|a) \equiv d \ln \sigma / d \ln \rho$. Using Eq. (12), we find at MLL accuracy

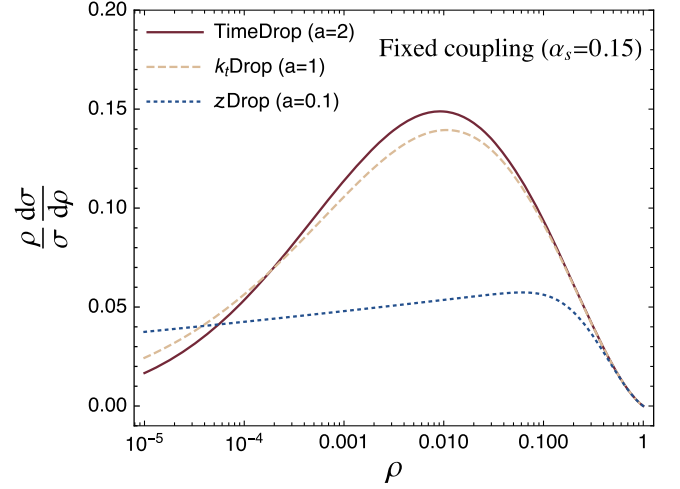


FIG. 3. The tagged mass distribution for fixed coupling as given by Eq. (19) for $2 > a > 0$.

$$\mathcal{H}(2|a) = \frac{2\bar{\alpha}}{2-a} \int_{\kappa}^{\kappa^{a/2}} \frac{dx}{x} \left[1 - \frac{3}{4} \frac{x^{2-a}}{\kappa^{2-a}} \right] \times \exp \left\{ -\frac{\bar{\alpha}}{a} \left[\ln^2 x + \frac{3}{2} (\ln x + 1 - x) \right] \right\}, \quad (19)$$

where in this case $\kappa \equiv \kappa^{(2)} = \rho = z(1-z)(\theta/R)^2$.

The result of numerically solving Eq. (19) is displayed in Fig. 3 for the most representative values of a , i.e., TimeDrop ($a=2$), k_t Drop ($a=1$) and z Drop ($a \approx 0$). We observe how the qualitative features generically discussed in the previous section are manifest: as a decreases the ρ distribution flattens while the peak shifts to larger values of ρ .

The case where $b = a = 2$ is of special interest. Using Eq. (18) we obtain

$$\frac{\rho d\sigma}{\sigma d\rho} = -\bar{\alpha} \left(\ln \rho + \frac{3}{4}(1-\rho) \right) \times \exp \left\{ -\bar{\alpha} \left[\frac{1}{2} \ln^2 \rho + \frac{3}{4} (\ln \rho + 1 - \rho) \right] \right\}, \quad (20)$$

which remarkably reproduces the result for the plain mass at leading-logarithmic accuracy.

This is not surprising since to this level of accuracy the plain mass is determined by the hardest splitting.

B. Tagged z distribution ($b=0$)

As a second example, we consider the tagged z distribution, with $b=0$ so that $\kappa^{(0)} = z(1-z) \approx z$. Since we are now dealing with a potentially infrared unsafe but Sudakov safe [20] observable [see Eq. (10)], one has to beware. However, defining $\mathcal{H}(0|a) \equiv d \ln \sigma / d \ln z$, it is straightforward to derive

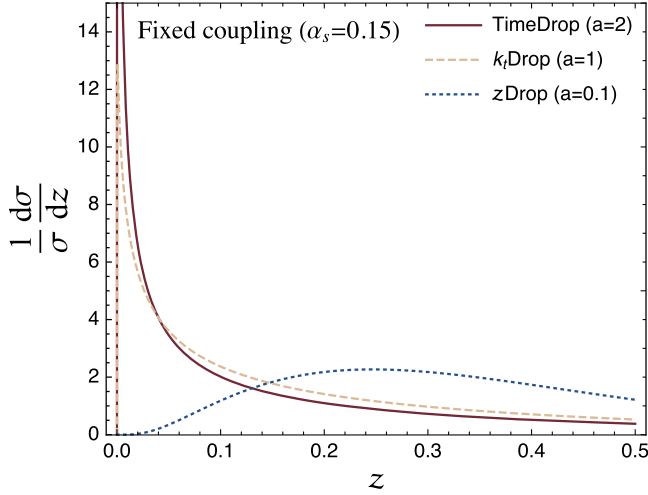


FIG. 4. The tagged z distribution for fixed coupling as given by Eq. (22) for $2 > a > 0$.

$$\frac{1}{\sigma} \frac{d\sigma}{dz} = P(z) \int_0^R \frac{d\theta}{\theta} \frac{d\alpha_s(k_t^2)}{\pi} \Delta(k|a). \quad (21)$$

After fixing the coupling and at MLLA, Eq. (21) transforms into

$$\frac{1}{\sigma} \frac{d\sigma}{dz} = \frac{2\bar{\alpha}}{za} \left(1 - \frac{3}{4}z\right) \int_0^z \frac{dx}{x} e^{-\frac{\bar{\alpha}}{a}[\ln^2 x + \frac{3}{2}(\ln x + 1 - x)]}. \quad (22)$$

The resulting tagged z distributions obtained by numerically solving Eq. (22) are displayed in Fig. 4 for $2 > a > 0$. The origin of the main features observed in Fig. 4 can be understood analytically by resorting to the DLA, where

$$\begin{aligned} \frac{1}{\sigma} \frac{d\sigma}{dz} &\approx \frac{2\bar{\alpha}}{za} \int_0^z \frac{dx}{x} e^{-\frac{\bar{\alpha}}{a} \ln^2 x} \\ &\approx \frac{1}{z} \sqrt{\frac{\bar{\alpha}\pi}{a}} \left[\operatorname{erf}\left(\sqrt{\frac{\bar{\alpha}}{a}} \ln z\right) - 1 \right]. \end{aligned} \quad (23)$$

This distribution is cut off at a characteristic value of z , namely

$$z_{\text{cut}} \approx e^{-\sqrt{a/\bar{\alpha}}}. \quad (24)$$

For $a \gg \bar{\alpha}$, this opens a wide range $z_{\text{cut}} < z < 0.5$ where the distribution falls off as z^{-1} and is modulated by $\sqrt{\bar{\alpha}/a}$. However, for $a \gtrsim 0$ and $z_{\text{cut}} \approx 1$, we find that

$$\frac{1}{\sigma} \frac{d\sigma}{dz} \sim \frac{1}{z} e^{-\frac{\bar{\alpha}}{a} \ln^2 z}; \quad (25)$$

i.e., the distribution grows slowly with z . These features are roughly reproduced in Fig. 4 where the dropoff for the k_t Drop case is clearly visible around $z \sim 0.02$.

In this context it is interesting to notice that the cutoff in z , Eq. (24), is dynamically generated and is a measure of $\sqrt{\alpha_s/a}$. This is quite different from SD (mMDT) grooming with $\beta = 0$ where the cutoff is simply given by the input to the algorithm. Although the distribution is modulated by the same ratio, dynamical grooming opens up the possibility to probe the splitting function down to low z .

IV. MONTE CARLO STUDIES AND RESILIENCE TO NONPERTURBATIVE EFFECTS

In this section, we complement our analytical studies by using PYTHIA8 [28] to simulate dijet events in proton-proton collisions at $\sqrt{s} = 13$ TeV. For each event, particles are clustered into anti- k_T jets [9] with $R = 0.8$ and reclustered with Cambridge/Aachen using FastJet 3.1 [29]. The analysis is performed on jets with $p_T > 450$ GeV/c. Further, the sensitivity to nonperturbative phenomena such as the underlying event (multiparton interactions and initial state radiation) and hadronization is explored.

We plot the kinematics of the tagged emissions on the primary Lund plane for the three main choices of a in Eq. (1), corresponding to TimeDrop ($a = 2$), k_t Drop ($a = 1$) and z Drop ($a \approx 0$), in Fig. 5. It is clear from these figures that the condition on the hardest branch in each of these three cases corresponds to suppressing the phase space at large formation times (alternatively, small virtualities), small k_t 's or small momentum fractions z , respectively. It is important to point out that there are no *sharp*

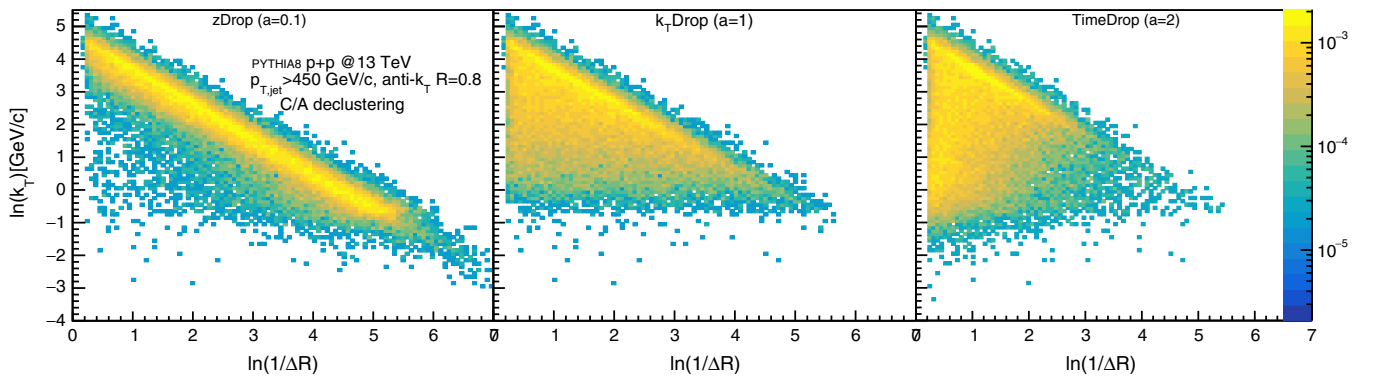


FIG. 5. Primary Lund planes for the tagged emissions.

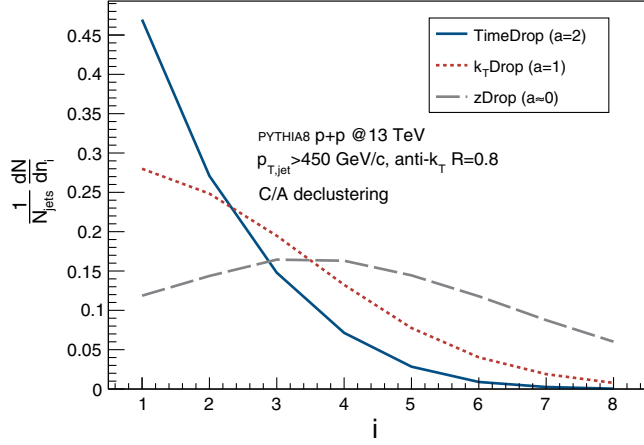


FIG. 6. Distribution of the number of the hardest branching on the primary Lund plane.

cuts in the kinematical plane, in contrast to other existing grooming algorithms, such as trimming, filtering, pruning and SoftDrop. This remarkable feature arises due to the fact that the hardest emission, which can be thought of as a proxy of the realistic jet scale, is fluctuating on a jet-by-jet basis. Nevertheless, a dynamical cut is generated which can be estimated by solving $\Delta(\kappa|a) = 1/2$. Up to DLA we find

$$\kappa = e^{-\sqrt{\frac{a \ln 2}{\alpha}}}, \quad (26)$$

or, in terms of $k_t = z\theta/R$,

$$\ln k_{t,\text{cut}}(\theta) = (1-a) \ln \frac{R}{\theta} - \sqrt{\frac{a \ln 2}{\alpha}}. \quad (27)$$

This defines a straight line boundary that is clear from the MC simulations of the Lund planes in Fig. 8. Parametrically, the point where the critical line crosses the y -axis can be estimated within the fixed coupling approximation to scale as

$$\ln k_{t,\text{cut}}(R) \sim -\sqrt{a}. \quad (28)$$

As an another illustration of the dynamical grooming, we plot in Fig. 6 the distribution of the number i of the tagged, hardest branching. Although infrared-collinear (IRC) unsafe it is useful to investigate the location of the tagged branching in the C/A sequence. The larger the power a , the more narrow and peaked around $i = 1$, i.e., the first C/A declustering step, the distribution is. This is quite natural since $a = \infty$ corresponds to an angular-ordered Sudakov form factor. In the opposite limit, $a \rightarrow 0$, the distribution widens and peaks around $i \gg 1$. More precisely, the average i in each grooming setting is ≈ 2 (TimeDrop), 3 (k_t Drop) and 5 (z Drop).

The ρ -distribution is presented in Fig. 7. As anticipated, both TimeDrop and k_t Drop exhibit a plain masslike shape while SoftDrop and z Drop deliver an almost flat

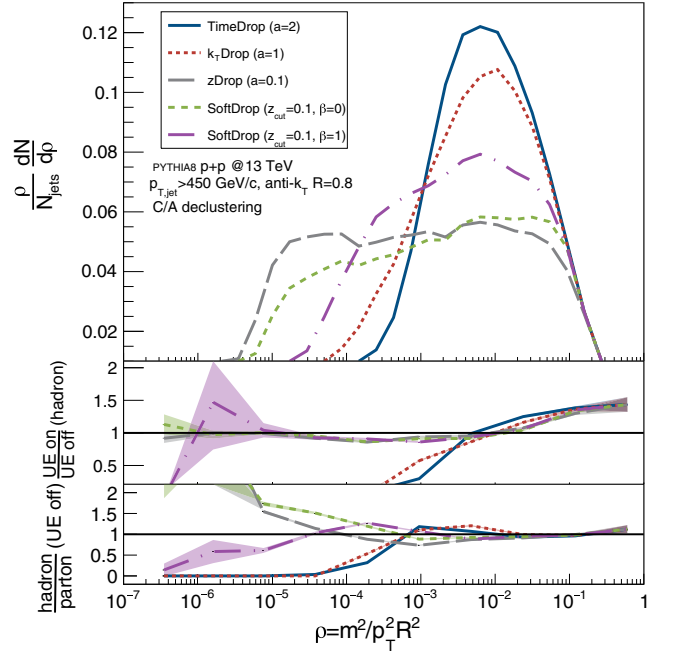


FIG. 7. Top: The tagged ρ -distribution for different choices of a in Eq. (1) and SoftDrop. Middle: Ratio of the distributions with and without the underlying event at hadron level. Bottom: Ratio of the distributions at hadronic and partonic level without the underlying event.

distribution. The sensitivity to nonperturbative physics is alike for every scenario, especially for values of $\rho > 10^{-3}$. It is worth noticing that while z Drop has a fantastic robustness against the underlying event, k_t Drop outperforms when considering hadronization. Therefore, we expect that a compromise to reduce the sensitivity to both mechanisms simultaneously could be achieved by using intermediate values of a i.e., $1 > a > 0.1$. This possibility together with an extended performance study of the method will be presented in an upcoming publication.

In the top panel of Fig. 8, the z distribution of the tagged splitting for different grooming procedures at partonic level and without the underlying event is displayed. For completeness, we show the results obtained with SoftDrop for $z_{\text{cut}} = 0.1$ and two different choices of $\beta = 0, 1$. We find an excellent agreement between the qualitative features of the analytic estimate for the dynamical grooming family as shown in Fig. 4 and the more realistic scenario provided by full-fledged Monte Carlo simulations. It is worth noticing the different behavior between SoftDrop ($\beta = 0$) and z Drop even though they use the same variable for tagging, i.e., the momentum sharing. Regarding the impact of nonperturbative effects, in the central and bottom panels we evaluate the role of the underlying event and hadronization, respectively. We would like to highlight the resilience of k_t Drop to hadronization effects, an imprint of its effectiveness on selecting the most perturbative splitting. For the other cases, an overall similar performance to SoftDrop is found.

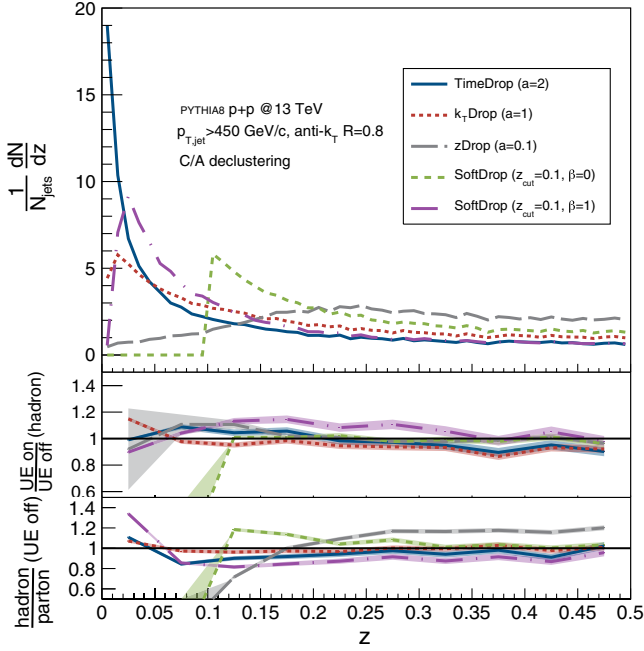


FIG. 8. Top: The tagged z distribution for different choices of a in Eq. (1) and SoftDrop. Middle: Ratio of the distributions with and without the underlying event at hadron level. Bottom: Ratio of the distributions at hadronic and partonic level without the underlying event.

V. CONCLUSIONS

In this work we have proposed a new set of jet substructure observables defined by the hardest splitting in a C/A reclustered jet. We explore three representative definitions of hardness in terms of formation time (TimeDrop), relative transverse momentum of the splitting (k_t Drop) and momentum sharing (z Drop). For a tagged splitting in the shower by either of these three choices, its kinematics serve to compute any observable such as its mass or momentum sharing fraction. Other observables, such as the groomed angular distribution, can be derived in a completely analogous fashion.

We have developed an analytical framework that gives a good qualitative understanding of the features seen in full Monte Carlo simulations. The key object in these calculations is the Sudakov form factor Eq. (3) that vetoes all primary emissions in the full angular range of the jet. While many contemporary grooming procedures involve two parameters, our approach relies on the intrinsically generated jet scale whose proxy is the hardness defined via the continuous parameter a . The amount of grooming is somehow related to how different the observable is from the variable that defines the hardness, i.e., by comparing b to a , where $a=b$ results in plain distributions. We found also that, in contrast to SD/mMDT where z_{cut} is a parameter of the grooming, a similar cutoff scale is naturally generated by the strong QCD dynamics, $z_{\text{cut}} \sim e^{-\sqrt{a}/a}$.

The observables discussed in this work are IRC safe except for the z distribution which turns out to be Sudakov safe.

So far we have only investigated observables exploiting the tagged hardest splitting inside a jet. In addition to the remarkable features of the analytic distributions, our Monte Carlo studies indicate that these observables are quite resilient to nonperturbative effects, including both hadronization and the underlying event, for a large part of the distributions. We find it particularly interesting to note that even with relatively mild grooming, $b \lesssim a$, the mass distribution is robust in the region of its peak (this is also the case for the k_t distribution). We propose to study such observables experimentally as they represent, perhaps, the closest realization of perturbative parton dynamics in fully fledged jet observables.

In this work we have deliberately avoided studying in more detail the grooming mode, where branches that violate the ordering set by the hardest branching would be removed, leading to modifications of the jet kinematics. This procedure naturally lends itself to an interpretation of removing radiation sensitive to the total color charge of the jet. It could easily be implemented in a recursive fashion along all the primary and secondary branches/planes of the jet. This will be studied in more detail in an upcoming paper.

ACKNOWLEDGMENTS

We are grateful to Jesse Thaler for a careful reading of the manuscript and useful comments. Y. M.-T. and K. T. acknowledge the ExtreMe Matter Institute, GSI, for hospitality and the participants of the Rapid Reaction Task Force Workshop “The space-time structure of jet quenching: theory and experiment” for illuminating discussions. The work of Y. M.-T. and A. S.-O. was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contract No. DE-SC0012704, and by Laboratory Directed Research and Development (LDRD) funds from Brookhaven Science Associates. K. T. is supported by a Starting Grant from Trond Mohn Foundation (Grant No. BFS2018REK01) and the University of Bergen.

APPENDIX A: COLLINEAR UNSAFE TAGGING ($a=0$)

The case $a=0$ has to be treated with special care, since it is formally collinear unsafe. In this case it is also unclear whether the hardest emission actually will be one of the primary emissions. Leaving this complication aside for the moment, in this case we would write Eq. (9) as

$$\ln \Delta(\kappa|0) = - \int_0^1 dz' \int_0^{R^2} \frac{d\theta'^2}{\theta'^2} \frac{\alpha_s(z'^2 p^2 \theta'^2)}{2\pi} P(z') \times \Theta(z'^2 > z^2) \Theta(z'^2 p_1^2 \theta'^2 > Q_0^2), \quad (\text{A1})$$

where we have demanded that $k'_i > Q_0$ and Q_0 be a nonperturbative scale.

For fixed coupling and in DLA, this reads

$$\begin{aligned} \ln \Delta(\kappa|0) &= -\bar{\alpha} \int_z^1 dz' P(z') \ln \frac{z' Q}{Q_0} \\ &\approx -\bar{\alpha} \ln \frac{1}{z} \ln \frac{z Q^2}{Q_0^2}. \end{aligned} \quad (\text{A2})$$

As a concrete example, let us consider the mass distribution in this case. The expression becomes

$$\frac{m^2}{\sigma} \frac{d\sigma}{dm^2} = \frac{1}{\mathcal{N}} \int_{\max[\frac{m^2}{Q_0^2}, \frac{Q_0^2}{m^2}]}^1 dz \frac{\alpha_s(zm^2)}{2\pi} P(z) \Delta(\kappa|0), \quad (\text{A3})$$

where the second condition on the integral comes about by demanding that $k'_i > Q_0^2$. We notice a strong shape sensitivity to the ratio Q^2/Q_0^2 . The normalization factor appears from the unitarity condition and reads

$$\begin{aligned} \mathcal{N} &= 1 - \exp \left[- \int_0^{R^2} \frac{d\theta^2}{\theta^2} \int_0^1 dz \frac{\alpha_s(k_t^2)}{2\pi} P(z) \Theta(k_t^2 > Q_0^2) \right] \\ &\approx 1 - \exp \left(-\bar{\alpha} \ln^2 \frac{Q}{Q_0} \right), \end{aligned}$$

where the last line was obtained in DLA for fixed coupling.

APPENDIX B: TAGGED k_t DISTRIBUTION ($b=1$)

For completeness, we provide the distribution corresponding to $b=1$, that is, the relative momentum of

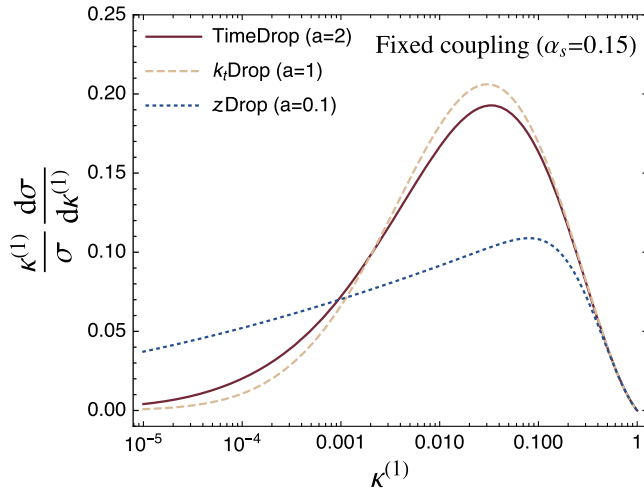


FIG. 9. The tagged $\kappa^{(1)}$ distribution for fixed coupling as given by Eq. (B1) for $2 > a > 0$.

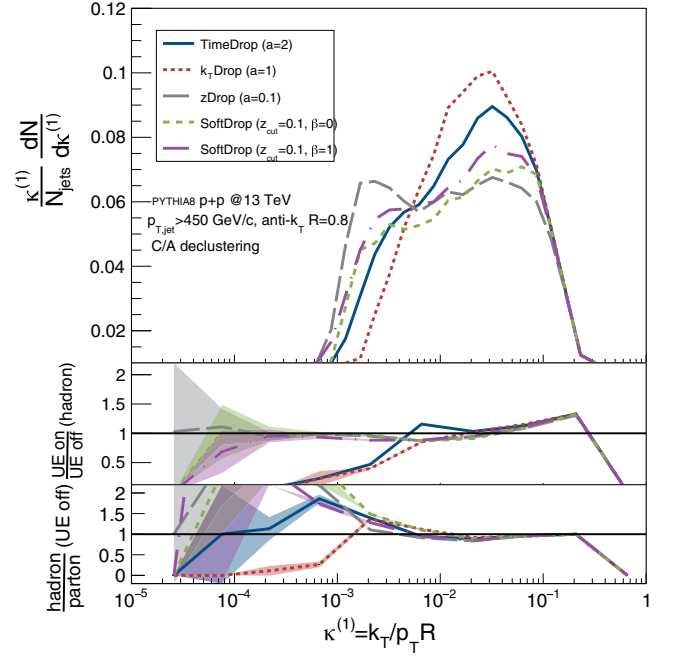


FIG. 10. Top: The tagged k_t distribution for different choices of a in Eq. (1) and SoftDrop. Middle: Ratio of the distributions with and without the underlying event at hadron level. Bottom: Ratio of the distributions at hadronic and partonic level without the underlying event.

the splitting, $\kappa^{(1)} = k_t/Q$. The tagged $\kappa^{(1)}$ distribution is given by

$$\begin{aligned} \mathcal{H}(1|a) &= \frac{2\bar{\alpha}}{1-a} \int_{\kappa}^{\kappa^a} \frac{dx}{x} \left[1 - \frac{3x^{1-a}}{4\kappa^{1-a}} \right] \\ &\times \exp \left\{ -\frac{\bar{\alpha}}{a} \left[\ln^2 x + \frac{3}{2} (\ln x + 1 - x) \right] \right\}. \end{aligned} \quad (\text{B1})$$

After performing the integral in Eq. (B1) numerically, we obtain the curves displayed in Fig. 9. In this case, the ordering in the peak position is reverted; i.e., it is located at larger values for k_t Drop as compared to TimeDrop. This fact follows naturally from Eq. (16).

Regarding the Monte Carlo results shown in Fig. 10, a good qualitative agreement with the corresponding analytic formulas is found together with a similar performance among the different scenarios with respect to the impact of the underlying event and hadronization.

APPENDIX C: GLUON-INITIATED JET

All through this manuscript we have considered quark-initiated jets as can be deduced from the splitting function given in Eq. (8). The study of gluon-initiated jets at MLLA amounts to replacing the color factor C_F by C_A and introducing the $g \rightarrow g + g$ splitting function

$$P(z) = P_{gg}(z) \simeq C_A \left[\frac{2}{z} - \frac{11}{6} \right]. \quad (\text{C1})$$

The Sudakov factor then becomes

$$\ln \Delta(\kappa|a) \simeq -\frac{\bar{\alpha}}{a} \left[\ln^2 \kappa + \frac{11}{6} (\ln \kappa + 1 - \kappa) \right], \quad (\text{C2})$$

and the rest of the calculation follows analogously as for the quark.

To understand the interplay between the color representation of the initiating parton and the values of a and b in Eq. (12) we have generated a distribution equivalent to the one displayed in the middle panel of Fig. 2 including the gluonic case. The result for $\mathcal{H}(b|1)$ is shown in Fig. 11. The gluon distributions are peaked at larger values of κ compared to the quark case as a result of their bigger color charge that translates into an enhanced Sudakov suppression, a well-known result [30]. Note that this transition point occurs at smaller values of κ when $b \sim 0$. A more thorough investigation of the quark/gluon-discriminating power of dynamical grooming will be presented in an upcoming publication.

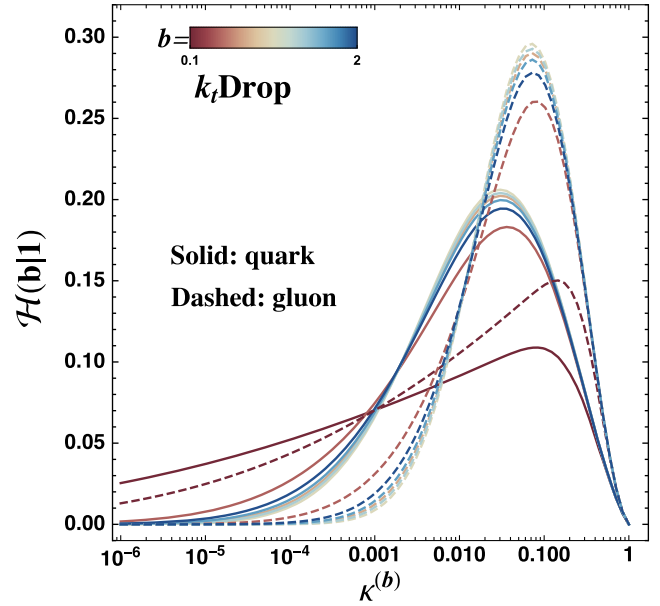


FIG. 11. The normalized distribution $\mathcal{H}(b|1)$ for a range of b values in Eq. (12) for fixed coupling at MLLA considering quark (solid) and gluon-initiated (dashed) jets.

-
- [1] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Phys. Rev. Lett.* **100**, 242001 (2008).
- [2] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **12** (2015) 055.
- [3] M. Dasgupta and G. P. Salam, *Phys. Lett. B* **512**, 323 (2001).
- [4] A. J. Larkoski, I. Moult, and B. Nachman, [arXiv:1709.04464](https://arxiv.org/abs/1709.04464).
- [5] R. Kogler *et al.*, *Rev. Mod. Phys.* **91**, 045003 (2019).
- [6] S. Marzani, G. Soyez, and M. Spannowsky, *Lect. Notes Phys.* **958** (2019).
- [7] S. D. Ellis and D. E. Soper, *Phys. Rev. D* **48**, 3160 (1993).
- [8] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *J. High Energy Phys.* **08** (1997) 001.
- [9] M. Cacciari, G. P. Salam, and G. Soyez, *J. High Energy Phys.* **04** (2008) 063.
- [10] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, *J. High Energy Phys.* **09** (2013) 029.
- [11] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, *Eur. Phys. J. C* **73**, 2623 (2013).
- [12] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, *J. High Energy Phys.* **05** (2014) 146.
- [13] F. A. Dreyer, L. Necib, G. Soyez, and J. Thaler, *J. High Energy Phys.* **06** (2018) 093.
- [14] D. Krohn, J. Thaler, and L.-T. Wang, *J. High Energy Phys.* **02** (2010) 084.
- [15] A. M. Sirunyan *et al.* (CMS Collaboration), *Phys. Rev. Lett.* **120**, 142302 (2018).
- [16] K. Kauder (STAR Collaboration), *Nucl. Phys.* **A967**, 516 (2017).
- [17] S. Acharya *et al.* (ALICE Collaboration), *J. High Energy Phys.* **10** (2018) 139.
- [18] S. Acharya *et al.* (ALICE Collaboration), [arXiv:1905.02512](https://arxiv.org/abs/1905.02512).
- [19] H. A. Andrews *et al.*, [arXiv:1808.03689](https://arxiv.org/abs/1808.03689).
- [20] A. J. Larkoski, S. Marzani, and J. Thaler, *Phys. Rev. D* **91**, 111501 (2015).
- [21] S. Marzani, L. Schunk, and G. Soyez, *Eur. Phys. J. C* **78**, 96 (2018).
- [22] Z.-B. Kang, K. Lee, X. Liu, and F. Ringer, *J. High Energy Phys.* **10** (2018) 137.
- [23] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Phys. Rev. D* **80**, 051501 (2009).
- [24] B. Andersson, G. Gustafson, L. Lonnblad, and U. Pettersson, *Z. Phys. C* **43**, 625 (1989).
- [25] F. A. Dreyer, G. P. Salam, and G. Soyez, *J. High Energy Phys.* **12** (2018) 064.
- [26] P. Nason, *J. High Energy Phys.* **11** (2004) 040.
- [27] A. J. Larkoski and J. Thaler, *J. High Energy Phys.* **09** (2013) 137.
- [28] T. Sjostrand, S. Mrenna, and P. Z. Skands, *Comput. Phys. Commun.* **178**, 852 (2008).
- [29] M. Cacciari, G. P. Salam, and G. Soyez, *Eur. Phys. J. C* **72**, 1896 (2012).
- [30] A. J. Larkoski, G. P. Salam, and J. Thaler, *J. High Energy Phys.* **06** (2013) 108.