

Investigating top tagging with Y_m -Splitter and N-subjettiness

Mrinal Dasgupta and Jack Helliwell

*Lancaster-Manchester-Sheffield Consortium for Fundamental Physics,
School of Physics & Astronomy, University of Manchester, Manchester M13 9PL, U.K.*

E-mail: mrinal.dasgupta@manchester.ac.uk,
jack.helliwell@manchester.ac.uk

ABSTRACT: We study top-tagging from an analytical QCD perspective focussing on the role of two key steps therein: a step to find three-pronged substructure and a step that places constraints on radiation. For the former we use a recently introduced modification of Y-Splitter, known as Y_m -Splitter, and for the latter we use the well-known N-subjettiness variable. We derive resummed results for this combination of variables for both signal jets and background jets, also including pre-grooming of the jet. Our results give new insight into the performance of top tagging tools in particular with regard to the role of the distinct steps involved.

KEYWORDS: QCD Phenomenology

ARXIV EPRINT: [2108.09317](https://arxiv.org/abs/2108.09317)

Contents

1	Introduction	1
2	Tagger definitions	4
3	Monte-Carlo study	5
4	Y_m-Splitter splitter with a τ_{32} cut: QCD jets	7
4.1	Leading-order result	8
4.2	Resummed results	9
4.2.1	The small τ limit	9
4.3	Finite τ corrections	11
4.3.1	Differential distribution in τ	12
4.3.2	Cumulative distribution	14
4.3.3	Pre-grooming with Soft Drop	16
4.3.4	Numerical implementation and parton shower studies	18
5	Signal jets	21
5.1	Jet mass distribution for top jets	22
5.2	Top jets with Y_m -Splitter	23
5.3	Y_m -Splitter with grooming for signal jets	26
5.4	Y_m -Splitter with τ_{32} and grooming for signal jets	27
6	Exploiting jet mass cuts	30
7	Conclusions	32
A	Finite τ approximations at fixed order	36
B	Direct calculation of the cumulative background distribution	37

1 Introduction

The past decade has seen the emergence of jet substructure as one of the key areas of LHC phenomenology [1–11]. The main impetus driving this emergence has come from studies involving boosted heavy particles, after pioneering work in the context of Higgs searches revealed the clear potential of substructure based analyses [3]. Following early studies, a rapid proliferation of tools and methods followed, mainly aiming at enhancing the discriminating power of substructure methods in various contexts including methods for tagging of two-pronged (e.g. W/Z/H decays) and three-pronged (i.e. top decay) jet

substructure [2–6, 12–14, 14–17] as well as for quark-gluon discrimination [18–20]. For in-depth reviews of these topics and further references we refer the reader to the review articles refs. [11, 21, 22].

While the rapid development of substructure methods often resulted in novel powerful techniques, many of which are currently still in use, some key questions also emerged about the robustness of the methods being employed. Such questions were concerned, for instance, with the accuracy to which Monte Carlo event generators provide a reliable description of substructure observables and about the dependence of tagger performance on poorly understood physics aspects like non-perturbative effects in QCD. This led to a parallel effort to better understand jet substructure as relevant to tagging and grooming of jets originating from boosted particles, from the first principles of QCD theory [23–35]. As a consequence it was possible to identify flaws in existing tools [23, 24], design superior tools which remove some of the main flaws thus identified [23, 26], and shed light on the factors that influence performance including the role of non-perturbative effects [23, 29–31, 36, 37].¹

The more recent advent of machine-learning (ML) tools to study jets has also yielded impressive performance gains with ML based taggers shown to often significantly outperform standard (“QCD-based”) tagging algorithms [37–47]. Nevertheless the questions raised for earlier tagging methods in terms of exclusive reliance on parton showers to study performance and the issue of performance gains originating in non-perturbative effects remain in the ML case, and are indeed potentially re-enforced. Here one can mention studies that have investigated the resilience of Lund-plane based ML [36, 37] against non-perturbative effects, finding that eliminating the non-perturbative region results in a marked decrease in performance. Furthermore new research on parton showers has revealed flaws in the *perturbative structure* of dipole showers including a failure to reproduce the QCD double emission matrix-element for soft emissions strongly ordered in angle [48, 49], in principle a crucial regime for meaningfully describing jet substructure.

Given all of the above, it therefore remains of importance to continue to develop the program of understanding jet substructure taggers via perturbative QCD. While much success has been obtained in analytic understanding of the impact of taggers and groomers on signal and background for two-pronged decays, there is a more limited understanding of top tagging which is a somewhat more complicated problem owing in part to the coloured parton initiating the signal jet. In terms of tools, various methods for finding three-pronged jet substructure have been introduced in the literature including the early ATLAS top tagger [16] based on Y-splitter, as well as the CMS top tagger [14, 14, 15] conceptually related to the mMDT/Soft Drop procedure [23, 26]. Other widely used methods for top tagging include the Johns Hopkins top tagger [12] and the HEP top tagger [13], shower deconstruction [50] and template tagging [51].

Amongst methods aiming at constraining radiation around three hard prongs, the N-subjettiness ratio τ_{32} [52] and Energy Correlation Function ratios [53] have been actively

¹In this context a notable feature that has often emerged in a variety of contexts is the presence of a trade-off between performance and resilience to non-perturbative effects [30, 36, 37].

studied. Combinations of these tools with grooming have also been studied and exploited in experimental analyses. For example refs. [54, 55] makes use of trimmed jets with a top tagging procedure involving a combination of Y-splitter and the N-subjettiness ratios τ_{32} and τ_{21} while ref. [56] reports, amongst other studies, combinations of the CMS top tagger with a τ_{32} cut. Such combinations are similar in essence to the combinations we shall study in the present article, though various details differ.

A first analytical study of the impact of prong-finding methods for top tagging supplemented with grooming, in the high p_T limit i.e. with p_T in the TeV range, was carried out in ref. [34]. This work included the study of IRC safe extensions of the IRC unsafe CMS top tagger as well as studying an adaptation of the Y-splitter method, Y_m -Splitter.

In this article we extend the work of ref. [34] by combining prong-finding with Y_m -Splitter, with an additional radiation constraint coming from a τ_{32} cut. Further, we account for the impact of pre-grooming with Soft Drop (SD) and mMDT. We begin with a set of Monte Carlo studies that motivate the use of this particular combination of methods as well as indicate optimal values for the τ_{32} cut, $\tau \sim 0.2$. Next we obtain resummed analytic results for QCD background jets for Y_m -Splitter with the τ_{32} cut, in the small τ limit. These results are obtained in a modified leading-logarithmic approximation where other than capturing all leading-logarithmic (LL) terms, one also retains important classes of next-to-leading-logarithmic (NLL) terms such as those from hard-collinear emission. Following the treatment of ref. [35] we then extend our results to include finite τ effects which are in general non-negligible even for our typical value of $\tau \sim 0.2$. We study both the un-groomed case as well as consider pre-grooming with Soft Drop ($\beta = 2$) and the mMDT. Despite this rather complex combination of methods leading to a highly non-trivial observable, we find that our results are in broad agreement with those from parton shower studies, with remaining moderate differences consistent with the expected size of omitted (beyond LL) terms.

Next we study signal jets on a similar footing. We start with a simple situation with only a mass window cut and compare the resulting Sudakov form factor to results from Pythia, finding excellent agreement. This step is useful in order to test the validity of our simplifying assumptions about radiation in a top initiated jet. We then extend our studies to include Y_m -Splitter in addition to the mass window cut, also accounting for pre-grooming using both mMDT and SD ($\beta = 2$). Our results here improve upon previous work by accounting for a previously neglected situation where one of the prongs found by Y_m -Splitter can be a soft gluon rather than one of the top decay products. The inclusion of this correction term brings our results for the signal into substantially better agreement with Pythia simulations, than was seen in previous studies [34]. We then account for the impact of τ cut in the signal case. Although our treatment of finite τ corrections for the signal is not as accurate as the corresponding treatment for the QCD background, we obtain a good description of the τ dependence of the result especially for the un-groomed case and for pre-grooming with SD($\beta = 2$).

The layout of this paper is as follows: we start in section 2 by recalling the definitions of the Y_m -Splitter tagger, N-subjettiness including our choice of axes, and Soft Drop grooming. In section 3 we report results from initial Monte Carlo studies which help lay

the ground for our subsequent analytical investigation. In section 4 we carry out our calculations for Y_m -Splitter with a τ_{32} cut for QCD background jets. Here we discuss in detail the small τ limit as well as accounting for finite τ effects, studying both the differential distribution and the cumulant. We close this section by including grooming with both mMDT and SD ($\beta = 2$) and comparing our results to those from Herwig and Pythia showers. Section 5 is devoted to signal jets where we first study the effect of a mass window cut alone followed by studies of Y_m -Splitter including grooming and finally the inclusion of a τ_{32} cut. Section 6 discusses our analytical results in terms of the understanding gained for the performance of the Y_m -Splitter, τ_{32} and grooming combinations in terms of the interplay between the τ and mass window cuts, and reports further comparisons to parton showers. Our conclusions are summarised in section 7.

2 Tagger definitions

The primary step involved in top-tagging is the identification of three-pronged jet substructure that characterises top-decay. There are various methods that have been suggested in the literature for the identification of three-pronged substructure within a fat jet, some of which have also been used for phenomenology. Examples of prong finding methods include the early CMS and ATLAS top taggers [14, 57, 58] and Y_m -Splitter, an adaptation of Y-splitter introduced in ref. [34] which we shall use for our analytical studies here. Additionally jet shape variables such as N-subjettiness [52], which we also use here, are known to be powerful methods that quantify the N-pronged nature of a jet through placing constraints on radiation from N identified prongs within a fat jet. Techniques combining prong-finding methods with jet shape variables are also known to give rise to important performance gains, have been used in experimental studies [54–56, 59] and motivate our desire to better understand such combinations. We define in more detail below all the specific methods that we use in this article.

1. Y_m -Splitter

The Y_m -Splitter method for top tagging [34] takes a jet clustered with the gen- k_t ($p = \frac{1}{2}$) algorithm (referred to as gen- k_t from here on) and performs the following steps:

- (a) Undo the last clustering, to give two sub-jets, both of which are examined for the condition $p_{t,i} > \zeta p_{t,\text{jet}}$. If either sub-jet fails the ζ condition, the jet is rejected.
- (b) Check which sub-jet produces the larger gen- k_t distance when de-clustered, and undo the last clustering of this sub-jet. Check whether the resulting sub-jets from this de-clustering pass the ζ condition. If either the de-clustering or the ζ condition fail, the jet is rejected.
- (c) Find the pairwise masses of the three final sub-jets, and require that $\min(m_{12}, m_{13}, m_{23}) > m_{\text{min}}$. If this condition is not met, the jet is rejected.

2. N-subjettiness

We will use the N-subjettiness ratio variable $\tau_{32} = \frac{\tau_3}{\tau_2}$, where

$$\tau_N^{(\beta)} = \frac{1}{p_{t,\text{jet}} R^\beta} \sum_{i \in \text{jet}} p_{t,i} \min \left((\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \dots, (\Delta R_{N,i})^\beta \right), \quad (2.1)$$

with the sum over i running over the jet constituents, $\Delta R_{ij} = \sqrt{(\Delta y_{ij})^2 + (\Delta \phi_{ij})^2}$, and the N partition axes are labelled $1 \cdots N$. There are various options for defining the partition axes, for instance finding the axes which minimise τ_N (optimal axes). Throughout this work we use $\beta = 2$ as this facilitates our analytical studies and we make use of the gen- k_t axes with $p = 1/2$. These axes are obtained by clustering the jet with the gen- $k_t(p = \frac{1}{2})$ algorithm and identifying the axes with the N exclusive sub-jets resulting from $N - 1$ de-clusterings. For τ_2 with $\beta = 2$, these have been shown to be very close to the optimal axes [35]. For τ_3 these axes are exactly the three prongs returned by Y_m -Splitter which is helpful in facilitating the resummation of the tagged fraction of events.

3. Soft Drop

Soft Drop takes a jet, re-clusters it with the Cambridge/Aachen (C/A) algorithm [60, 61] and performs the following steps:

- (a) Undo the last clustering, to give two sub-jets.
- (b) Examine the lower p_T sub-jet for the condition $p_{t,i} > z_{\text{cut}}(\frac{\Delta R}{R})^\beta (p_{t,i} + p_{t,j})$.
- (c) If this condition is not met this sub-jet is removed from the jet and the groomer goes back to step (a). If it is met, the groomer stops and this is the final jet.

Throughout this work we set $z_{\text{cut}} = \zeta$, the Y_m -Splitter parameter.

3 Monte-Carlo study

In this section we investigate the performance of various tagging procedures based around the N-subjettiness variable τ_{32} as well as how they are impacted by hadronisation, ISR, and MPI. The tagging procedures considered all have the restriction that the jet mass is between 160 GeV and 225 GeV, corresponding to a window around the top mass, and are studied as a function of the cut on τ_{32} . After examining N-subjettiness cuts with and without pre-grooming we combine these cuts with the Y_m -Splitter method which we again investigate with and without pre-grooming. Two pre-grooming options are considered, SD ($\beta = 2$) and mMDT (equivalent to SD with $\beta = 0$).

We start by generating 1 million $t\bar{t}$ and $q\bar{q}$ events with Pythia.² ISR, MPI and hadronisation were initially deactivated, and a generation cut of $p_t > 1600$ GeV was applied. Jets were clustered with the Cambridge/Aachen algorithm with $R = 1$ and $p_{t,\text{min}} = 2$ TeV using Fastjet 3 [62], as was the case for the studies in ref. [34].

Where jets are groomed we use $z_{\text{cut}} = 0.05$. τ_{32} is calculated using the N-subjettiness fastjet contrib [63], and where Y_m -Splitter is used we choose $m_{\text{min}} = 50$ GeV and $\zeta = 0.05$. This information is then used to construct the tagged fraction of events and the signal to square-root-background as a function of a cut on τ_{32} . The same procedure is used both with only hadronisation, and then hadronisation, ISR and MPI activated to assess their impact.

²We have studied the impact of including gluon jets and found that in the p_t range under consideration they do not significantly modify the distributions presented.

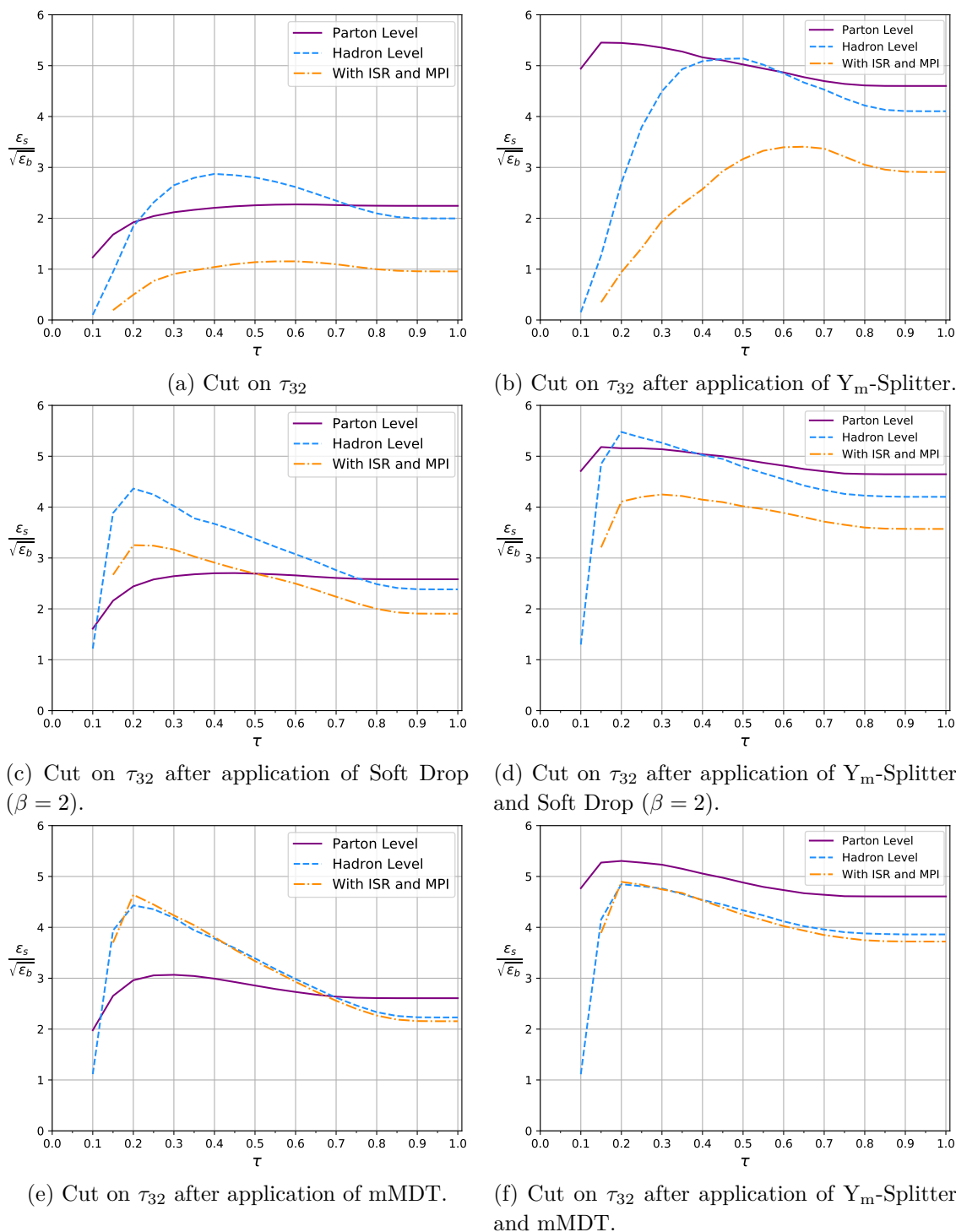


Figure 1. Plots showing the signal to square-root background of the four variants of the tagging procedure at parton level, hadron level, and with ISR and MPI activated at hadron level.

To discuss the features that emerge from our Monte Carlo studies let us first examine the top row of figure 1, i.e. figures 1(a) and 1(b), which show the signal significance as a function of the τ_{32} cut, τ , without any grooming and without Y_m -Splitter on the left and with Y_m -Splitter on the right. It is clear that in the absence of a grooming step ISR and MPI significantly damage performance in each case, although the inclusion of Y_m -Splitter results in a higher signal significance after all effects are considered.

Next we come to the plots involving the application of grooming i.e. figures 1(c) and 1(d) for SD ($\beta = 2$) pre-grooming and figures 1(e) and 1(f) in the bottom row for the mMDT. From these one notes that grooming, especially with mMDT, is an effective method to significantly mitigate ISR and MPI. When combining grooming with Y_m -Splitter we observe that both hadronisation and ISR+MPI are significantly reduced, resulting in high performance with an optimal value of $\tau \sim 0.2$ emerging for mMDT pre-grooming and $\tau \sim 0.3$ for SD ($\beta = 2$). The best performance, i.e. highest signal significance, comes with mMDT pre-grooming and Y_m -Splitter applied in addition to the τ cut, as shown in figure 1(f). This combination is also more resilient to ISR and all non-perturbative effects at the same time. In contrast although pre-grooming jets and cutting on τ_{32} without Y_m -Splitter (see figure 1(e)) gives good performance at hadron level, the discrepancy with the parton level result indicates that the performance of this procedure cannot necessarily be understood from perturbative QCD arguments alone and may be more susceptible to mis-modelling of non-perturbative effects in parton showers.³

In summary, applying Y_m -Splitter to pre-groomed jets with cuts on τ_{32} and the jet mass is a high performing method for tagging hadronically decaying high- p_T top quarks.⁴ The performance is also well described by parton level predictions and is therefore reasonably robust against effects which are less well theoretically understood in this context. These observations provide some of the main motivation for detailed theoretical studies using perturbative QCD, which will be the subject of the next two sections.

4 Y_m -Splitter splitter with a τ_{32} cut: QCD jets

We start by examining the impact of a τ_{32} cut on QCD jets after applying Y_m -Splitter. Analytical studies for Y_m -Splitter as applied to top-tagging, with and without pre-grooming, have already been carried out in ref. [34]. These studies derived results for the jet mass distribution and consequently the efficiency for QCD jets tagged with Y_m -Splitter using the technique of QCD resummation. Resummation is required in order to address the multi-scale nature of the problem. Crucially the highly boosted limit implies that the invariant jet mass $m^2 \ll p_T^2$, with $m^2 \sim m_t^2$ and p_T values in the TeV range, which leads to large logarithms in $\rho = m^2/R^2 p_T^2$. A good description of the jet-mass distribution

³A possible reason for this might be that a pure τ_{32} cut is not IRC safe and is instead only Sudakov safe [64, 65] while the application of Y_m -Splitter prior to the subjeettiness cut prevents τ_2 from vanishing, resulting in an IRC safe quantity.

⁴We find that, for comparable signal significance, these methods appear, in the high p_T region, to outperform the dense neural net and boosted decision tree used by ATLAS in [66], although it should be noted that the two studies are perhaps not equivalent, as no attempt was made here to examine detector effects, which were included in the ATLAS study.

then requires resummation of the logarithms in ρ . Additionally for Y_m -Splitter we have $\rho_{\min} = m_{\min}^2/p_T^2 R^2 \ll 1$ and a further small scale ζp_T , the minimum energy of an emission that passes the ζ condition, with $\zeta \ll 1$. Large logarithms are then expected and do arise in ρ , ρ_{\min} , ζ and in ρ_{\min}/ρ . In ref. [34] a modified leading logarithmic resummation was performed which included all double-logarithmic terms and a subset of single-logarithmic terms such as those arising from hard-collinear emissions. The logarithms that are most crucial to resum are those in the smallest parameters ρ and ρ_{\min} . Typical values of $\zeta \sim 0.05$ and $\rho_{\min}/\rho \sim m_W^2/m_{\text{top}}^2$ are larger and hence we only aim to retain logarithms in these parameters at leading double-logarithmic accuracy.

Here, relative to previous work [34] we shall additionally include the τ_{32} cut, considering the possibility that τ_{32} is not small. In doing so we shall follow closely the treatment of ref. [35] for resummation of jet mass with a τ_{21} cut.

4.1 Leading-order result

We start with the leading-order result, computed in the soft and collinear approximation which yields the leading logarithmic terms. For Y_m -Splitter this starts at order α_s^2 for QCD jets, since one requires at least two emissions within the jet (i.e. at least three partons) in order to be accepted by Y_m -Splitter. Since for three partons τ_3 vanishes, a cut requiring $\tau_{32} < \tau$ is trivially satisfied. Therefore the leading-order result is unchanged from the pure Y_m -Splitter case of ref. [34]. For the case of a quark initiated jet and in the abelian C_F^2 channel it is given by⁵

$$\frac{1}{\sigma} \left(\frac{d\sigma}{d\rho} \right)^{\text{LO,soft-collinear}} = \bar{\alpha}^2 \int \frac{dz_1}{z_1} \frac{dz_2}{z_2} \frac{d\theta_1^2}{\theta_1^2} \frac{d\theta_2^2}{\theta_2^2} \times \Theta(\theta_2^2 < \theta_1^2 < 1) \delta(\rho - \max(z_1\theta_1^2, z_2\theta_2^2)) \times \Theta(z_1 > \zeta_{\text{cut}}) \Theta(z_2 > \zeta_{\text{cut}}) \Theta(\min(z_2\theta_2^2, z_1z_2\theta_1^2) > \rho_{\min}), \quad (4.1)$$

where we defined $\bar{\alpha} = \frac{C_F \alpha_s}{\pi}$, taking for definiteness the case of a quark initiated jet. In deriving the above result we have taken a strongly-ordered in angle configuration with $\theta_2 \ll \theta_1$, made a leading logarithmic approximation that the jet mass is dominated by the emission that makes the larger contribution, and imposed the tagger conditions by requiring both emissions to pass the ζ_{cut} and implemented the ρ_{\min} condition in the strongly-ordered limit where $\theta_{12} \sim \theta_1$.

We then obtain:

$$\frac{\rho}{\sigma} \left(\frac{d\sigma}{d\rho} \right)^{\text{LO,soft-collinear}} \stackrel{\rho_{\min} < \zeta_{\text{cut}}}{=} \bar{\alpha}^2 \ln^2 \frac{1}{\zeta_{\text{cut}}} \ln \frac{\rho}{\rho_{\min}}, \quad (4.2)$$

$$\stackrel{\rho_{\min} > \zeta_{\text{cut}}}{=} \bar{\alpha}^2 \ln^2 \frac{\rho}{\rho_{\min}} \left(\frac{3}{2} \ln \frac{1}{\zeta_{\text{cut}}} - \frac{1}{2} \ln \frac{\rho}{\rho_{\min}} \right).$$

A similar result is obtained for the $C_F C_A$ colour factor while in the $C_F T_R n_f$ channel the result is one logarithm down due to the lack of a soft enhancement in the p_{qg} splitting function. For future convenience we note that the leading-order result can also be expressed

⁵We shall define all our angles to correspond to the actual angles rescaled by the jet radius R .

in terms of the highest-mass emission ρ_a and the next-highest-mass emission ρ_b . Written in these terms we have

$$\begin{aligned} \frac{1}{\sigma} \left(\frac{d\sigma}{d\rho} \right)^{\text{LO,soft-collinear}} &= \bar{\alpha}^2 \int \frac{dz_a}{z_a} \frac{dz_b}{z_b} \frac{d\rho_a}{\rho_a} \frac{d\rho_b}{\rho_b} \delta(\rho - \rho_a) \\ &\times \Theta(\rho_a > \rho_b) \Theta(z_a > \zeta_{\text{cut}}) \Theta(z_b > \zeta_{\text{cut}}) \\ &\times \Theta(\min\{\rho_b, z_a z_b \max(\theta_a^2, \theta_b^2)\} > \rho_{\text{min}}), \end{aligned} \tag{4.3}$$

where in the ρ_{min} condition we used strong angular ordering to replace θ_{ab}^2 by $\max(\theta_a^2, \theta_b^2)$. Finally, we note that beyond double logarithmic accuracy a more precise result at order α_s^2 can be achieved by considering three collinear partons within a jet without imposing strong ordering between the partons. Such configurations are described by triple-collinear splitting functions and calculations implementing the triple-collinear result were included in the studies of Y_m -Splitter carried out in ref. [34].

4.2 Resummed results

Now we turn to the resummed result. We first consider the case where $\tau_{32} < \tau \ll 1$. Then we shall lift the requirement that $\tau \ll 1$ i.e. we shall account for finite τ effects.

4.2.1 The small τ limit

For the case of Y_m -Splitter one considers, as in ref. [34], two real emissions that pass the tagger cuts accompanied by an ensemble of soft and collinear emissions which are constrained to set a smaller gen- k_t distance (i.e. mass) than either of the two leading emissions. This constraint on real emissions produces a Sudakov form factor. In the current case the emissions are additionally constrained by the τ cut. Here we shall derive the Sudakov form factor at leading-logarithmic (LL) accuracy, capturing all double-logarithmic terms including those in τ and running coupling effects, and also include some important single logarithmic effects such as accounting for hard-collinear radiation.

For the two emissions accounted for at leading-order, eq. (4.2), we shall again label ρ_a as the emission that sets the larger mass and ρ_b the smaller mass. Consider first all subsequent *primary* emissions, i.e. emissions from the hard parton initiating the jet. These emissions must not give rise to larger mass (gen- k_t) values than the first two emissions de-clustered by Y_m -Splitter *and* they must set a value of $\tau_{32} < \tau$. Recall that the contribution of an emission i to τ_N is given by $z_i \min(\theta_{i1}^2, \dots, \theta_{iN}^2)$. As was the case for τ_2 [28], the limit of strong angular-ordering ensures that, for emissions coming from a leg lying along one of the N-subjettiness axes, the smallest of the θ_{ia} angles is either the angle between the emission and its emitter, or can be approximated by this angle to LL accuracy. For a primary emission this implies that the contribution to τ_3 , $\tau_{3i} = z_i \theta_i^2$ where z_i is the energy fraction and θ_i is the angle of the emission w.r.t. the hard initial parton. The value of τ_2 on the other hand is dominated, to LL accuracy, by the second highest mass emission ρ_b , due to the strong ordering in masses relevant at LL accuracy.⁶ The condition on primary

⁶For going beyond the small τ limit and including finite τ effects we shall, in the next subsection, lift this requirement of strong ordering in masses as was done in ref. [35].

emissions then reads:

$$\Theta\left(\frac{\sum_{i=3}^{\infty} z_i \theta_i^2}{\rho_b} < \tau\right) \prod_{i=3}^{\infty} \Theta\left(z_i \theta_i^2 < \rho_b\right). \quad (4.4)$$

The first step function reflects the condition on τ_{32} while the second condition reflects the constraint on mass which gives the primary emission Sudakov form factor for Y_m -Splitter in ref. [34], i.e. that none of the emissions i have a gen- k_t distance larger than ρ_b by assumption. Since $\tau < 1$ the second condition is automatically satisfied and the condition on primary emissions is just given by the stronger constraint $\Theta\left(\sum_{i=3}^{\infty} z_i \theta_i^2 < \rho_b \tau\right)$. The primary emission Sudakov factor then arises from a veto on any emissions violating this condition. More precisely it takes the form $S = e^{-R^{(\text{primary})}}$ with the ‘‘radiator’’ $R^{(\text{primary})}$ given by

$$R^{(\text{primary})}(\tau, \rho_b) = \frac{C_R}{2\pi} \int \alpha_s(z^2 \theta^2 R^2 p_T^2) p(z) dz \frac{d\theta^2}{\theta^2} \Theta(z\theta^2 > \rho_b \tau), \quad (4.5)$$

where C_R is a colour factor that depends on the identity of the initiating jet i.e. C_F for a quark and C_A for a gluon jet, and $p(z)$ is the QCD splitting function describing collinear emission from a quark ($p(z) = p_{gq}(z)$) or gluon ($p(z) = p_{gg}(z)$). For the argument of the running coupling we have used the k_t of the emission (in terms of z and θ) as required in the soft and collinear limit.

As well as vetoing primary emissions from the parton initiating the jet, the overall Sudakov factor must also account for a veto on secondary emissions which would set a value of τ_{32} larger than τ from either of the two emissions included in the leading order pre-factor. In the case of soft secondary emissions the angle of emission θ_i is limited by angular-ordering to be less than the angle of the parent θ_a or θ_b . Apart from this constraint, for emissions off parton a we have the same constraint as for primary emissions and hence we obtain, for emissions off parton a :

$$R^{(\text{secondary},a)}(\tau, \rho_a, \rho_b) = \frac{C_A}{2\pi} \int \alpha_s(z^2 z_a^2 \theta^2 R^2 p_T^2) p_{gg}(z) dz \frac{d\theta^2}{\theta^2} \Theta(z z_a \theta^2 > \rho_b \tau) \Theta(\theta^2 < \theta_a^2), \quad (4.6)$$

where we note that z represents the energy fraction of parton a ’s energy carried by the soft secondary emission. We also note that for secondary emissions, the gen- k_t distance, entering the veto condition above, differs from the mass even in the soft limit, involving one less factor of z_a . A similar equation gives the result for emissions off parton b , with the obvious replacement of z_a and θ_a by z_b and θ_b .

The overall result can be written as a Sudakov form factor weighting the leading-order, order α_s^2 result which serves as a pre-factor. For simplicity if one retains just the leading-logarithmic expression for the pre-factor reported in eq. (4.3), we can obtain the resummed result by inserting the factor $S = e^{-R}$ in the integrand in eq. (4.3) where $R = R^{(\text{primary})} + R^{(\text{secondary},a)} + R^{(\text{secondary},b)}$. While our results include the full running coupling and hard-collinear effects we report below a simplified result for the Sudakov factor S in the limit of a fixed coupling and retaining only the soft-collinear behaviour i.e. replacing $p(z)$ and $p_{gg}(z)$ by the soft limit expression $2/z$:

$$S^{(\text{fixed-coupling, soft})} = \exp\left[-\frac{C_R \alpha_s}{2\pi} \ln^2 \frac{1}{\tau \rho_b} - \frac{C_A \alpha_s}{2\pi} \ln^2 \frac{\rho_a}{\tau \rho_b} - \frac{C_A \alpha_s}{2\pi} \ln^2 \frac{1}{\tau}\right], \quad (4.7)$$

where the term involving $\ln^2 1/\rho\tau_b$ comes from primary emissions, the term involving $\ln^2 \frac{\rho_a}{\tau\rho_b}$ comes from vetoing emissions from emission a and finally the suppression involving just $\ln^2 1/\tau$ comes from vetoing emissions from emission b . The difference between primary and secondary emissions arises entirely from angular-ordering and the ensuing limitation on emission angle we mentioned previously. Although the logarithms present in S are written above in terms of ρ_a and ρ_b , these will eventually be related to logarithms of ρ , ρ_{\min} and ratios thereof once the integrals in the pre-factor are carried out. In the limit $\tau \rightarrow 1$ of the above result we obtain the pure Y_m -Splitter result of ref. [34].

A couple of further remarks are in order concerning the result in eq. (4.7). First of all the result captures leading double logarithms in τ in addition to the logarithms involved in the resummation of plain Y_m -Splitter [34]. Including hard-collinear emission via using the full splitting functions, rather than just their soft limit, and using the running coupling helps to improve the result beyond double-logarithmic accuracy. The result indicates that the effect of the N-subjettiness cut is to produce an extra suppression relative to the case of Y_m -Splitter [34] just by changing the scale ρ_b to the smaller scale $\tau\rho_b$ and the extra secondary suppression factor we get from emissions off parton b . This suppression of the background is of course desirable but choosing a small τ value potentially also suppresses the signal, which in this case is a coloured particle namely the top quark. Also as is well-known from several prior applications [52, 63] and additionally emerges in the Monte Carlo studies we reported in section 3, optimal values of τ do not necessarily satisfy $\tau \ll 1$, so that finite τ effects generally need to be considered [35] in addition to the resummation of logarithms of τ . The inclusion of finite τ corrections is thus the topic of the next subsection.

4.3 Finite τ corrections

To obtain an insight into the role of the τ_{32} cut in a phenomenological context, one has to address values of $\tau \sim 1$. From the viewpoint of resummation this has implications identical to those first pointed out in the τ_{21} case [35]. The small τ limit resummation of the previous subsection is designed to fully capture double logarithmic terms of the form $\alpha_s^n L^{2n}$ where, for power counting purposes, we use the symbol L^{2n} to denote double logarithms in any of $\ln \rho$, $\ln \rho_{\min}$, $\ln \frac{\rho}{\rho_{\min}}$, $\ln \tau$ or any combination of them. From the fixed-coupling Sudakov form factor, eq. (4.7), written in terms of ρ_a, ρ_b and τ we note that we obtain terms that are single logarithmic in jet masses (and jet mass ratios) but double logarithmic overall due to the role of $\ln \tau$ i.e. terms of the form $\alpha_s \ln \rho_b \ln \tau$ and $\alpha_s \ln \frac{\rho_a}{\rho_b} \ln \tau$. Beyond the small τ limit we need to account for such terms beyond just their $\ln \tau$ dependence i.e. obtain the full function f_τ that multiplies single logarithms in jet mass. However, given that single logarithms in jet mass ratios i.e. $\alpha_s \ln \rho_b/\rho_a$ are smaller, we do not attempt to obtain the finite τ corrections for such terms which is substantially more involved and accordingly retain their small τ form only. In the Sudakov form factor with inclusion of running coupling, we therefore wish to control terms of the form $\alpha_s^n L_\rho^n f_n(\tau)$ (where L_ρ generically denotes logarithms in jet masses but not ratios), while the small τ resummation accounts only for terms that approximate $f_n(\tau)$ by its leading small τ behaviour $\sim \ln^n \tau$.

While in the small τ limit resummation of the previous subsection we assumed that emissions were strongly ordered in terms of their contribution to the jet mass (in addition

to strong ordering in angle), in order to achieve resummation of terms $\alpha_s^n L_\rho^n$ with their accompanying τ dependence we no longer assume strong ordering in jet masses. In particular we assumed that τ_2 was set by a single emission b , and hence its value was taken to be ρ_b . Beyond the small τ limit, we must account for the fact that τ_2 receives a contribution from all emissions in the jet except emission a . Since we still desire terms that are at least single logarithmic in jet masses, we continue to assume that emissions are strongly ordered in angle. This approximation of emissions ordered in angle but not in mass is the same as was made in the case of τ_{21} studies for two-pronged decays [35], to obtain the finite τ correction to the small τ results [28].

We can then write

$$\tau_{32} = \frac{\rho - \rho_a - \rho_b}{\rho - \rho_a}, \tag{4.8}$$

where the numerator comes from τ_3 being given by the sum of jet masses contributed by all emissions except a and b , while the denominator is τ_2 which is given by the sum of jet masses contributed by all emissions except emission a . The sum of all emissions' contributions to jet mass, including those of a and b , just gives the total jet mass ρ .

4.3.1 Differential distribution in τ

We begin by presenting a result for the joint distribution in jet mass ρ and τ i.e. the quantity $\frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau}$, where τ is a set value of τ_{32} . To begin with we shall consider primary emissions only, since it is straightforward to account for secondary emissions in the final result.

We first write the result for the cross-section differential in both τ and ρ , which accounts for the two emissions a and b included in the leading-order formula but now accompanied by an infinite number of additional emissions which are strongly ordered in angle. The strong ordering in angle ensures that these emissions are emitted independently from the hard initial parton which leads to the standard factorised formula for any number of emissions:

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} &= \bar{\alpha}_s^2 \int_\zeta^1 \frac{dz_a}{z_a} \int_0^1 \frac{d\rho_a}{\rho_a} \int_\zeta^1 \frac{dz_b}{z_b} \int_0^{\rho_a} \frac{d\rho_b}{\rho_b} \Theta\left(\min\left(\rho_b, z_a z_b \max\left(\frac{\rho_a}{z_a}, \frac{\rho_b}{z_b}\right)\right) > \rho_{\min}\right) \\ &\times \exp\left[-\int_0^1 R'(\rho') \frac{d\rho'}{\rho'}\right] \sum_{p=1}^\infty \frac{1}{p!} \prod_{i=1}^p \int_0^{\rho_b} R'(\rho_i) \frac{d\rho_i}{\rho_i} \rho \delta\left(\rho - \rho_a - \rho_b - \sum_{i \neq a,b} \rho_i\right) \\ &\times \tau \delta\left(\tau - 1 + \frac{\rho_b}{\rho - \rho_a}\right). \end{aligned} \tag{4.9}$$

The above result is written using a fixed-coupling approximation for the emission of partons a and b though we shall account for the running of the coupling, with the k_t of those emissions, in the pre-factor for our final results. It involves considering p factorised real emissions, with a sum over all p , alongside a sum over all virtual corrections included via the exponential form factor. The factor R' , appearing in both real and virtual terms above, stems from the integral over the emission probability for a single emission in the soft and collinear limit, at a fixed mass ρ :

$$R'(\rho) = C_R \int_0^1 \frac{d\theta^2}{\theta^2} dz p(z) \frac{\alpha_s(z\theta p_T R)}{2\pi} \rho \delta(z\theta^2 - \rho) \stackrel{\text{f.c.}}{=} \frac{C_R \alpha_s}{\pi} \left(\ln \frac{1}{\rho} + B_i\right), \tag{4.10}$$

where the r.h.s. of the above equation gives the fixed-coupling result and we have replaced the splitting functions $p(z)$ by their soft piece $\propto 1/z$ and incorporated the effect of hard-collinear emissions by introduction of the B_i terms, corresponding to inclusion of the hard-collinear piece of the splitting function to our accuracy. For quark and gluon jets respectively $B_q = -3/4$ and $B_g = (-11C_A + 4n_f T_R)/(12C_A)$.

Integrating over ρ_b in eq. (4.9) using the delta function constraint involving τ allows us to set $\rho_b = (1 - \tau)(\rho - \rho_a)$ which then leads to the result, assuming $\tau < 1/2$,

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} &= \bar{\alpha}^2 \frac{\tau}{1-\tau} \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \Theta\left(\rho_a > \frac{1-\tau}{2-\tau}\rho\right) \Theta_{\rho_{\min}}(\rho_a, \tau, \rho_{\min}, z_a, z_b) \\ &\times \exp\left[-\int_0^1 R'(\rho') \frac{d\rho'}{\rho'}\right] \sum_{p=1}^{\infty} \frac{1}{p!} \prod_{i=1}^p \int_0^{\rho_b} R'(\rho_i) \frac{d\rho_i}{\rho_i} \rho \delta\left((\rho - \rho_a)\tau - \sum_{i \neq a,b} \rho_i\right), \end{aligned} \quad (4.11)$$

where we have used the shorthand notation $\Theta_{\rho_{\min}}$ to denote the ρ_{\min} condition and the ρ_b that occurs as an upper limit in the ρ_i integral is understood to be the value set by the delta function i.e $(\rho - \rho_a)(1 - \tau)$. The other step function constraint on ρ_a derives from the condition that $\rho_a > \rho_b$ and the value of ρ_b set by the delta function integral we have performed. Finally we observe that we are left to evaluate the multiple emission contribution where the sum over the ρ_i are constrained to be equal to $(\rho - \rho_a)\tau$. Additionally each emission i is constrained so that $\rho_i < \rho_b = (1 - \tau)(\rho - \rho_a)$, however for $\tau < 1/2$ this condition is automatically met if the stronger condition on the sum of ρ_i is satisfied. In what follows we restrict our attention to $\tau < 1/2$ as this region is sufficient given the optimal value of τ that emerged from the Monte Carlo studies in section 3. Finally we note that one can evaluate the multiple emission contribution on the second line of eq. 4.9 simply by using known results for the standard jet mass [67], since the constraint on multiple emissions is the same as for the plain jet mass ρ but with ρ replaced by $(\rho - \rho_a)\tau$. Hence without needing to perform any further explicit calculation one can write:

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} &\stackrel{\tau < 1/2}{=} \bar{\alpha}^2 \frac{1}{1-\tau} \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho - \rho_a} \\ &\times \Theta\left(\rho_a > \frac{1-\tau}{2-\tau}\rho\right) \Theta_{\rho_{\min}}(\rho_a, \tau, \rho_{\min}, z_a, z_b) \\ &\times R'((\rho - \rho_a)\tau) \frac{\exp[-R((\rho - \rho_a)\tau) - \gamma_E R'((\rho - \rho_a)\tau)]}{\Gamma[1 + R'((\rho - \rho_a)\tau)]}. \end{aligned} \quad (4.12)$$

The above result accounts for configurations where the three prongs tagged by Y_m -Splitter are the hard parton which initiates the jet along with two gluons emitted independently from it, however, our final result also contains configurations where a gluon emitted from the hard parton branches, with the resulting three particles corresponding to the three Y_m -Splitter prongs.

For the region $\tau > 1/2$ one could in principle follow the same method as outlined for the τ_{21} calculation in ref. [35], though given our immediate motivation we do not consider this region further in the present study. We could also have initially integrated over ρ_a

instead of ρ_b to obtain the result in the equivalent form

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} \stackrel{\tau < 1/2}{\equiv} \bar{\alpha}^2 \frac{1}{1-\tau} \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\frac{\rho}{2}} \frac{d\rho_b}{\rho_b} \frac{\rho}{\rho - \frac{\rho_b}{1-\tau}} \Theta_{\rho_{\min}}(\rho_b, \tau, \rho_{\min}, z_a, z_b) \\ \times R' \left(\rho_b \frac{\tau}{1-\tau} \right) \frac{\exp[-R(\rho_b \frac{\tau}{1-\tau}) - \gamma_E R'(\rho_b \frac{\tau}{1-\tau})]}{\Gamma[1 + R'(\rho_b \frac{\tau}{1-\tau})]}. \end{aligned} \quad (4.13)$$

A key feature of our results is the presence of an overall $1/(1-\tau)$ factor as was also the case in the τ_{21} result of ref. [35]. Taking the small τ limit of eq. (4.12) we should return to the small τ result we derived in the previous subsection, which is indeed the case up to subleading terms in the order $\bar{\alpha}^2$ pre-factor. To be more precise, in the previous subsection we had evaluated the pre-factor taking ρ_a to dominate the jet mass, by using the condition $\delta(\rho - \rho_a)\Theta(\rho_a > \rho_b)$, which correctly captures all double logarithmic terms in the pre-factor. If instead one uses the more accurate condition $\delta(\rho - \rho_a - \rho_b)\Theta(\rho_a > \rho_b)$, then after integration over ρ_b we obtain the same result as the small τ limit of eq. (4.12). Relative to the strong ordering of emissions ρ_a and ρ_b , using the exact jet mass conditions affects only single logarithmic terms $\bar{\alpha}^2 L^2$ in the pre-factor where L denotes logarithms in ρ/ρ_{\min} or ζ . Such terms are two logarithms below the leading $\bar{\alpha}^2 L^4$ terms in the pre-factor and only involve more modest logarithms than those in the jet mass. We can therefore consider such terms as negligible and hence the small τ limit of eq. (4.12) is equivalent to the result of the previous subsection. For an explicit calculation demonstrating the argument above, we refer the reader to appendix A.

Beyond the small τ limit the most crucial feature of the result is the overall $1/(1-\tau)$ factor in eq. (4.12). While there is additionally a τ dependence in the step functions in eq. (4.12), that is again responsible for introducing τ dependent terms of order $\bar{\alpha}^2 L^2$ in the pre-factor, and hence can be neglected to our accuracy as illustrated in appendix A. In what follows we shall use the freedom to set τ to zero in the step function conditions to obtain an analytic form for the cumulative distribution.

4.3.2 Cumulative distribution

It is of direct interest to also obtain the result for the differential distribution in the jet mass with a cut on τ_{32} , $\tau_{32} < \tau$ rather than fixing a value for τ_{32} as required for the double differential distribution above. In order to do this one can integrate the differential distribution, eq. (4.12), between 0 and τ . Setting τ to zero in the step functions of the pre-factor, which has only a sub-leading impact as discussed before, we can write

$$\begin{aligned} \rho \frac{d\sigma}{d\rho} \Big|_{\tau_{32} < \tau} \stackrel{\tau < 1/2}{\equiv} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho - \rho_a} \Theta \left(\rho_a > \frac{\rho}{2} \right) \Theta_{\rho_{\min}}(\rho_a, \rho_{\min}, z_a, z_b) \\ \times \int_0^{\tau} \frac{d\tau'}{\tau'(1-\tau')} R'((\rho - \rho_a)\tau') \frac{\exp[-R((\rho - \rho_a)\tau') - \gamma_E R'((\rho - \rho_a)\tau')]}{\Gamma[1 + R'((\rho - \rho_a)\tau')]}, \end{aligned} \quad (4.14)$$

which corresponds to integrating the distribution up to some maximum value τ for τ_{32} . To single-logarithmic accuracy, we can expand the radiator about some point τ_0 to write:

$$R((\rho - \rho_a)\tau') \simeq R((\rho - \rho_a)\tau_0) - R'((\rho - \rho_a)\tau_0) \ln \left(\frac{\tau'}{\tau_0} \right) + \mathcal{O}(R''), \quad (4.15)$$

where $R'(x) = -\frac{\partial R}{\partial \ln x}$. With τ_0 chosen such that in the small τ limit τ_0 is of order τ , and given that the integral is dominated by values $\tau' \sim \tau$, terms of order R'' and beyond can be neglected as they are beyond single-logarithmic accuracy and we may replace τ' by τ_0 in the R' terms to obtain:

$$\begin{aligned} \rho \frac{d\sigma}{d\rho} \Big|_{\tau_{32} < \tau} &\stackrel{\tau < 1/2}{\equiv} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho - \rho_a} \Theta\left(\rho_a > \frac{\rho}{2}\right) \Theta_{\rho_{\min}}(\rho_a, \rho_{\min}, z_a, z_b) \\ &\times R'((\rho - \rho_a)\tau_0) \frac{\exp[-R((\rho - \rho_a)\tau_0) - \gamma_E R'((\rho - \rho_a)\tau_0)]}{\Gamma[1 + R'((\rho - \rho_a)\tau_0)]} \times I(R', \tau, \tau_0), \end{aligned} \quad (4.16)$$

where

$$I(R', \tau, \tau_0) = \int_0^{\tau} \frac{d\tau'}{\tau'(1 - \tau')} \exp\left[R'((\rho - \rho_a)\tau_0) \ln \frac{\tau'}{\tau_0}\right]. \quad (4.17)$$

Upon evaluating the integral over τ we obtain

$$I(R', \tau, \tau_0) = \left(\frac{\tau}{\tau_0}\right)^{R'} \frac{{}_2F_1(1, R', 1 + R', \tau)}{R'}, \quad R' \equiv R'((\rho - \rho_a)\tau_0). \quad (4.18)$$

We then have the result for the cumulative distribution given by eq. (4.16) with finite τ effects encoded in the Hypergeometric function of eq. (4.18) precisely as for the τ_{21} case [35]. The origin of the Hypergeometric factor is simply the extra overall factor of $1/(1 - \tau)$ in the finite τ differential distribution. Without this factor we simply obtain the usual result for the cumulative (integrated distribution) up to terms involving R'' beyond our accuracy, as long as $\tau_0 \sim \tau$. In what follows we shall simply choose $\tau_0 = \tau$ while noting that varying this choice by an $\mathcal{O}(1)$ factor will correspond to an effective resummation scale uncertainty on our results.

To obtain an alternate form of eq. (4.16) we could have integrated eq. (4.13) over τ instead. Again, as before, we can drop any τ dependence in the pre-factor other than the overall $\frac{1}{1 - \tau}$, which leads to a factor $\rho/(\rho - \rho_b)$, rather than $\rho/\left(\rho - \frac{\rho_b}{1 - \tau}\right)$. This again leads one to consider only the overall $1/(1 - \tau)$ factor together with the τ dependence in the exponent. Then integrating over τ using the same steps that gave eq. (4.16) we obtain

$$\begin{aligned} \rho \frac{d\sigma}{d\rho} \Big|_{\tau_{32} < \tau} &\stackrel{\tau < 1/2}{\equiv} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_{\rho_{\min}}^{\rho} \frac{d\rho_b}{\rho_b} \frac{\rho}{\rho - \rho_b} \Theta_{\rho_{\min}}(\rho_b, \rho_{\min}, z_a, z_b) \\ &\times {}_2F_1\left(1, R'\left(\rho_b \frac{\tau}{1 - \tau}\right), 1 + R'\left(\rho_b \frac{\tau}{1 - \tau}\right), \tau\right) \\ &\times \frac{\exp[-R(\rho_b \frac{\tau}{1 - \tau}) - \gamma_E R'(\rho_b \frac{\tau}{1 - \tau})]}{\Gamma[1 + R'(\rho_b \frac{\tau}{1 - \tau})]}, \end{aligned} \quad (4.19)$$

where we have again used the freedom to neglect factors of τ in the pre-factor which only introduce terms of order $\bar{\alpha}_s^2 L^2$ and set τ_0 to τ .

While so far we have worked with a fixed-coupling approximation in our pre-factor, we now introduce the running of the coupling for “emissions” a and b . In order to do so we replace the $\bar{\alpha}^2$ term with $\bar{\alpha}(z_a \rho_a p_T^2 R^2) \times \bar{\alpha}(z_b (\rho - \rho_a) p_T^2 R^2)$ inside the integral of eq. (4.16). This corresponds to using the k_t of each emission in the argument of the corresponding

coupling factor, with neglect of a factor $1 - \tau$ in the coupling associated to emission ρ_b , i.e. using $\rho_b = (\rho - \rho_a)$ instead of $(\rho - \rho_a)(1 - \tau)$. The $1 - \tau$ factor only results in sub-leading terms involving logarithms of $1 - \tau$ which we neglect, consistent with our general treatment of the pre-factor.

Finally to include secondary emissions we use the full radiator including the secondary emission terms i.e. replace

$$R((\rho - \rho_a)\tau) \rightarrow R^{(\text{primary})}((\rho - \rho_a)\tau) + R^{(\text{secondary,a})}((\rho - \rho_a)\tau, z_a, \theta_a^2) + R^{(\text{secondary,b})}((\rho - \rho_a)\tau, z_b, \theta_b^2), \quad (4.20)$$

where $\theta_a^2 = \frac{\rho_a}{z_a}$, $\theta_b^2 = \frac{\rho_b}{z_b}$ and $\rho_b = (\rho - \rho_a)$.

4.3.3 Pre-grooming with Soft Drop

It is known that the Y-Splitter and Y_m -Splitter methods need to be supplemented by some form of grooming in order to yield good performance for the signal significance (signal to square-root of background ratio) [29, 30, 34]. In ref. [30] it was found, in the context of W/Z/H tagging, that pre-grooming jets with Soft Drop was optimal in terms of increasing performance while minimising the sensitivity to non-perturbative effects. Furthermore, in the context of top-tagging there is another advantage to pre-grooming, namely that the pre-grooming procedure leads to a Sudakov form factor inherited from the groomer [30]. In other words for mMDT pre-grooming we obtain the mMDT Sudakov structure while for Soft Drop with non-zero β we obtain the Soft Drop Sudakov for both signal and background jets. Given that a modest rather than strong Sudakov suppression was found to be beneficial for signal significance in top-tagging [34], pre-grooming with mMDT which has only a single-logarithmic Sudakov form factor, followed by Y_m -Splitter, emerged as the most performant method as well as being resilient to non-perturbative effects.

Here we consider QCD jets pre-groomed with mMDT as well as Soft Drop for $\beta = 2$. In ref. [34] a result was obtained for the jet mass distribution with Soft Drop pre-grooming followed by the application of Y_m -Splitter i.e. without the additional τ cut involved here. As described in detail in ref. [34], three situations can arise: a) the largest gen- k_t emission, i.e. a in the present paper, stops the groomer, b) the next largest gen- k_t emission, i.e. b stops the groomer and c) another emission stops the groomer. For the first situation the result obtained for the primary emission radiator, with mMDT grooming, was shown to be of the form:

$$R^{(1), (\text{primary})}(\theta_a, \rho_b) = R_{\text{mMDT}}(\rho_b) + R_{\text{mMDT}}^{\text{angle}}(\theta_a, \rho_b). \quad (4.21)$$

This corresponds to the usual mMDT Sudakov at the scale ρ_b but modified by the addition of an extra piece, $R_{\text{mMDT}}^{\text{angle}}$ that arises because emissions with angle below θ_a are not examined by the groomer and hence need to be vetoed (if they have mass above ρ_b) even if they have $z < \zeta$. This extra contribution, at fixed-coupling and leading logarithmic accuracy, is given by [34]:

$$R_{\text{mMDT}}^{\text{angle}}(\theta_a, \rho_b) = \frac{C_R \alpha_s}{\pi} \int \frac{dz}{z} \frac{d\theta^2}{\theta^2} \Theta(z < \zeta) \Theta(z\theta^2 > \rho_b) \Theta(\theta_a^2 > \theta^2). \quad (4.22)$$

In case b), where emission b stops the tagger, one obtained instead just the standard mMDT result $R_{\text{mMDT}}(\rho_b)$, while for case c) where an emission other than a or b stops the tagger, there is a complete cancellation against virtual corrections and hence no contribution.

For our current work where we apply also a τ cut, situation a) yields the result reported in eq. (4.21) but now the mass scale ρ_b is replaced by $\tau(\rho - \rho_a)$ in both terms of eq. (4.21). In the case b) where emission b stops the tagger we now have to also account for the fact that while emissions with $z < \zeta$ and $\theta < \theta_b$ can never set a mass, or equivalently gen- k_t distance, above ρ_b , they can set a mass larger than $\tau(\rho - \rho_a)$. This is disallowed by the τ cut and hence such emissions have to be vetoed which leads to the appearance of a term $R_{\text{mMDT}}^{\text{angle}}(\theta_b, \tau(\rho - \rho_a))$, in addition to $R_{\text{mMDT}}(\tau(\rho - \rho_a))$, also in case b).

Taking into account hard-collinear emissions and the running of the coupling we can write our result in the form

$$\begin{aligned}
 R_{\text{groomed-mMDT}}^{(\text{primary})}((\rho - \rho_a)\tau, \theta_1) &= \int \frac{C_R \alpha_s(z\theta p_T)}{\pi} \left(\frac{1}{z} + B_i\right) dz \frac{d\theta^2}{\theta^2} \Theta(z\theta^2 > (\rho - \rho_a)\tau) \Theta(z > \zeta) \\
 &+ \int \frac{C_R \alpha_s(z\theta p_T)}{\pi} \left(\frac{1}{z} + B_i\right) dz \frac{d\theta^2}{\theta^2} \Theta(z\theta^2 > (\rho - \rho_a)\tau) \Theta(z < \zeta) \Theta(\theta^2 < \theta_1^2),
 \end{aligned} \tag{4.23}$$

where the first line is just the standard mMDT result [23], the second line is the extra R^{angle} contribution and $\theta_1 = \max(\theta_a, \theta_b)$ is the angle of the emission which stops the groomer. The basic form of the result is then that of the mMDT Sudakov evaluated at the scale $(\rho - \rho_a)\tau$, which corresponds to a single-logarithmic Sudakov suppression. In a fixed-coupling leading log approximation, the R^{angle} term can be written as

$$R^{\text{angle}}(\theta_1, (\rho - \rho_a)\tau) = \frac{C_R \alpha_s}{2\pi} \ln^2 \frac{\zeta \theta_1^2}{(\rho - \rho_a)\tau} \Theta(\theta_1^2 \zeta > (\rho - \rho_a)\tau), \tag{4.24}$$

where the logarithm involves a ratio of two small quantities similar to be behaviour obtained for secondary emission contributions. Overall therefore we retain the feature that pre-grooming with mMDT results in a reduced Sudakov suppression factor relative to the un-groomed case. The step function in eq. (4.24) switches off the R^{angle} contribution when $\theta_1^2 < \frac{(\rho - \rho_a)\tau}{\zeta}$ leading to the two regimes shown on the Lund diagrams in figure 2, where the R^{angle} piece is active only in figure 2(a) and is responsible for vetoing emissions in the region of phase space shown in blue. The standard soft drop Sudakov factor at the scale $(\rho - \rho_a)\tau$ is responsible for vetoing the region of phase space shown in red in figure 2.

One can also consider pre-grooming with Soft Drop. Identical considerations to the mMDT case apply, with the only difference being in the grooming condition i.e. for an emission to pass the grooming one needs $z > \zeta \theta^\beta$. We then obtain a result along similar lines to that for the mMDT above, but with the Soft Drop Sudakov (i.e. radiator) replacing that for the mMDT and a corresponding R^{angle} contribution whose fixed-coupling leading-log form is explicitly reported in ref. [34].

Secondary emissions are unaffected by grooming so the only change to the radiator, relative to the un-groomed case, arises from the primary emission term discussed above. The inclusion of finite τ effects is also unchanged relative to our previous discussions so

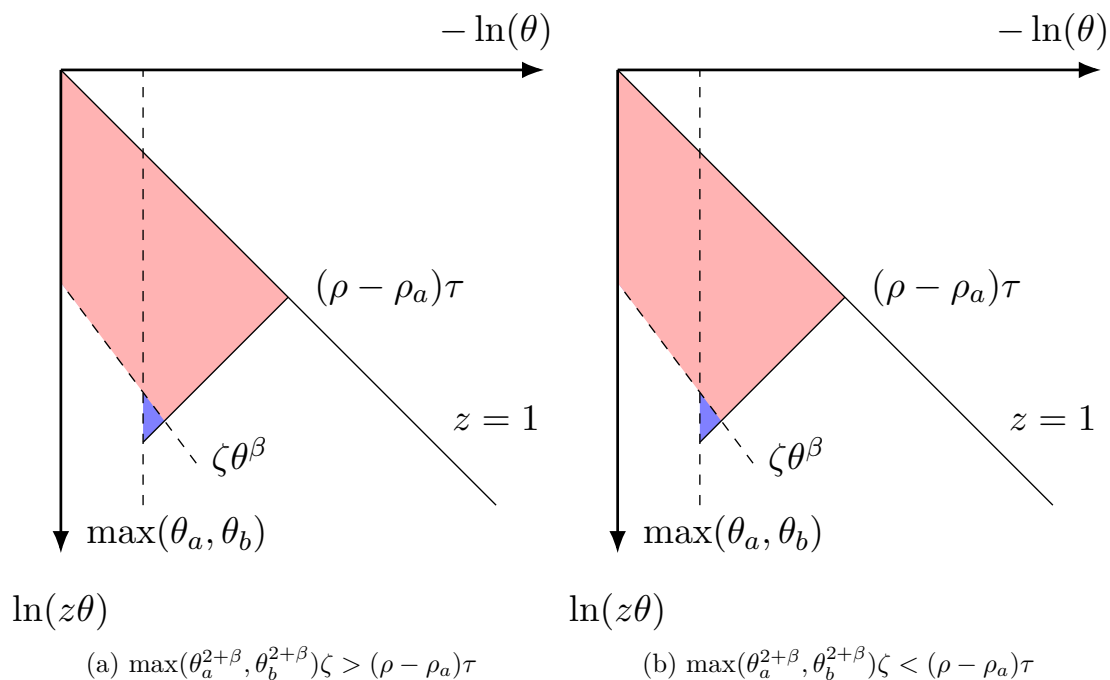


Figure 2. Lund diagrams showing the region of phase space vetoed for jets which are groomed with Soft Drop and tagged with Y_m -Splitter and a cut on τ_{32} .

that we still have the result eq. (4.12) for the differential distribution and eq. (4.16) for the cumulant but with the primary emission radiator replaced by that for the groomed case eq. (4.23) for mMDT and its analogue for Soft Drop.

4.3.4 Numerical implementation and parton shower studies

For the rest of this section we focus on quark initiated jets, as in the jet p_T range under consideration, these are the dominant background to top jets, though most of what follows could equally be applied to gluon initiated jets with minimal modifications. The form of eq. (4.19) is that of the leading $\mathcal{O}(\alpha_s^2)$ result multiplied by a factor accounting for further emissions. We now perform a type of matching to improve the accuracy with which we calculate this leading order pre-factor. While we have mentioned in section 4.1 that a more precise calculation of the leading order pre-factor based around the triple collinear splitting functions is possible, it was shown in [34] that the numerical difference between such a calculation and one using a product of $1 \rightarrow 2$ splitting functions, but the full phase-space, is slightly less than 10% for a jet mass of 175 GeV and $m_{\min} = 50$ GeV. Further to this, when a pair of collinear emissions are strongly ordered in angle, as we have considered them to be throughout this work, the appropriate matrix element is a product of $1 \rightarrow 2$ splitting functions. We therefore choose to match our resummed calculation on to a pre-factor calculated by taking the matrix element to be a product of $1 \rightarrow 2$ splitting functions but still using the full three-particle phase-space in the collinear limit. This particular matching procedure also potentially serves to bring the effects included in our calculations more in line with what is captured by the parton showers which we will compare our

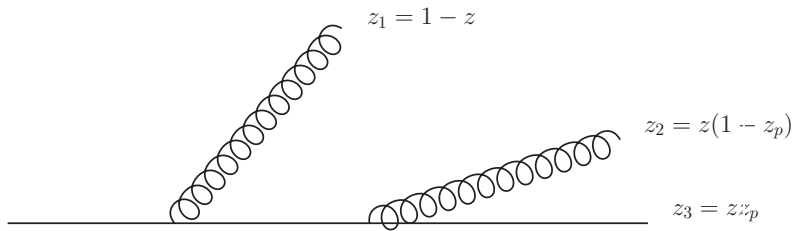


Figure 3. Diagram showing the parametrisation of the energy fraction variables used along with our labelling of the partons in the C_F^2 channel.

calculations to, as while these may be expected to contain elements of the phase-space, they do not include the full triple collinear splitting functions.

We now re-calculate the LO pre-factor at this higher level of accuracy, before showing how it is matched to the full resummation. As before, we use the C_F^2 channel for illustrative purposes, although our final results contain contributions from the $C_F C_A$ and $C_F n_f$ colour channels where similar modifications can be made to those listed below. In what follows, the parton initiating the jet is labelled as parton 3, with the emission at the widest angle to this parton labelled with 1 and the smaller angle emission labelled 2. So as to ensure that the variables appearing as the arguments of the factorised $1 \rightarrow 2$ splitting functions are defined appropriately, we work with the energy fraction variables z and z_p defined so that $z_1 = 1 - z$ and $z_2 = z(1 - z_p)$ as illustrated in figure 3. At this level of accuracy, the leading-order calculation in section 4.1 becomes

$$\begin{aligned} \frac{d\Sigma^{\text{LO}}(\tau, \rho)}{d\rho} &= \left(\frac{C_F \alpha_s}{2\pi} \right)^2 \int p_{gq}(z) p_{gq}(z_p) dz dz_p \Delta^{-1/2} \frac{d\theta_{12}^2}{\pi} \frac{d\theta_{13}^2}{\theta_{13}^2} \frac{d\theta_{23}^2}{\theta_{23}^2} \\ &\times \delta \left(\rho - \frac{s_{123}}{p_T^2 R^2} \right) \Theta_{\text{Y}_m\text{-Splitter}} \Theta_{\text{clust.}} \Theta(\theta_{13} > \theta_{23}) \end{aligned} \quad (4.25)$$

where the Gram determinant is given by

$$\Delta = 4\theta_{13}^2 \theta_{23}^2 - (\theta_{13}^2 - \theta_{12}^2 - \theta_{23}^2)^2, \quad (4.26)$$

and

$$\Theta_{\text{Y}_m\text{-Splitter}} = \Theta(\min(\rho_{12}, \rho_{13}, \rho_{23}) > \rho_{\text{min}}) \Theta(\min(z_1, z_2, z_3) > \zeta), \quad (4.27)$$

where $\rho_{ij} = z_i z_j \theta_{ij}^2$, encapsulating the conditions imposed by $\text{Y}_m\text{-Splitter}$ without approximating any particles as soft. Similarly, without approximating any particles as soft

$$\Theta_{\text{clust.}} = \sum_{i < j \neq k} \Theta(\theta_{ij} < \min(\theta_{ik}, \theta_{jk})) \Theta(\theta_{ij} < R) \Theta(\theta_{ij,k} < R). \quad (4.28)$$

Although the LO part of $\left. \frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \right|_{\tau_{32} < \tau}$ is now calculated without approximating any of the three LO partons as soft, we cannot simply substitute it in place of the $\mathcal{O}(\alpha_s^2)$ part of eq. (4.19). We must first specify how the quantities appearing in the Sudakov factor of eq. (4.19), which are defined in the soft and collinear limit, are related to the kinematic

variables appearing in our improved LO pre-factor (eq. 4.25). For the C_F^2 channel we make the following prescription:

$$\begin{aligned}
 \rho_b &= \min(\min(zz_p, (1-z))\theta_{13}^2, z \min(z_p, 1-z_p)\theta_{23}^2) \\
 k_{t1} &= \min(z, 1-z)\theta_{13}Rp_T & k_{t2} &= z \min(z_p, 1-z_p)\theta_{23}Rp_T \\
 \theta_1 &= \theta_{13} & \rho &= \frac{s_{123}}{p_T^2 R^2} + \sum_i \rho_i
 \end{aligned} \tag{4.29}$$

which we note that there is some freedom in choosing, the only constraint being that the correct result must be recovered in the soft and strongly-ordered limit.

Replacing the $\mathcal{O}(\alpha_s^2)$ part of equation (4.19) with eq. (4.25) and using the matching prescription given in eq. (4.29) we can write:

$$\begin{aligned}
 \rho \frac{d\Sigma^{\tau < \frac{1}{2}}(\tau)}{d\rho} &= \left(\frac{C_F}{2\pi}\right)^2 \int \alpha_s(k_{t1})\alpha_s(k_{t2})p_{gq}(z)p_{gq}(z_p)\Delta^{-1/2}\Theta(\theta_{13} > \theta_{23}) \\
 &\times \delta\left(\rho - \frac{s_{123}}{p_T^2 R^2}\right)\Theta_{\text{Y}_m\text{-Splitter}}\Theta(\theta_{13} > \theta_{23}) \\
 &\times {}_2F_1\left(1, R'\left(\rho_b \frac{\tau}{1-\tau}\right), 1 + R'\left(\rho_b \frac{\tau}{1-\tau}\right), \tau\right) \\
 &\times \frac{e^{-R(\rho_b \frac{\tau}{1-\tau}) - \gamma_E R'(\rho_b \frac{\tau}{1-\tau})}}{\Gamma[1 + R'(\rho_b \frac{\tau}{1-\tau})]} dz dz_p \frac{d\theta_{12}^2}{\pi} \frac{d\theta_{13}^2}{\theta_{13}^2} \frac{d\theta_{23}^2}{\theta_{23}^2}, \tag{4.30}
 \end{aligned}$$

where the quantities ρ_b , k_{t1} , and k_{t2} are as defined in eq. (4.29).⁷ Eq. (4.19) can be recovered from eq. (4.30) by replacing $s_{123} \rightarrow \rho_a + \rho_b$, neglecting the hard collinear part of the splitting functions, carrying out the θ_{12} integral (equivalent to an azimuthal integral) and changing phase-space variables back to ρ_a, ρ_b, z_a and z_b . For the sake of brevity, the above result is given only for the C_F^2 colour channel, however, our final results include the $C_F C_A$ and $C_F n_f$ colour channels, where a single gluon is emitted and then decays as opposed to the two independent emissions shown above. We also include secondary Sudakov factors in our final result exactly as before. Our results for pre-groomed jets are obtained by replacing the primary radiator with the groomed variant as discussed in section 4.3.3.

Eq. (4.30) is evaluated numerically using the Suave numerical integrator [68] interfaced to Mathematica [69]. As the cut on τ_{32} restricts emissions down to very low transverse momenta, we freeze the running coupling at $k_t = 1.5$ GeV to prevent divergences due to the Landau pole. The tagged background fraction is constructed from eq. (4.30) by integrating ρ over the mass window. This is shown in figure 4 along with the same quantity derived from parton shower simulations using both Pythia and Herwig [70] for three variations

⁷In deriving equation (4.30) all emissions are considered to contribute to the jet mass as shown in eq. (4.29). As well as allowing us to capture the function of τ multiplying single logarithms in mass scales, this also generates τ dependant terms which are beyond our accuracy. These terms are removed, as discussed in section 4.3, by neglecting the τ dependence in the pre-factor beyond the leading $1/1-\tau$ term which leads to the hypergeometric function in eq. (4.30). Specifically, we have set τ to zero inside the delta function, which would otherwise be written as $\delta\left(\rho - \frac{s_{123}}{p_T^2} + \mathcal{O}(\tau)\right)$.

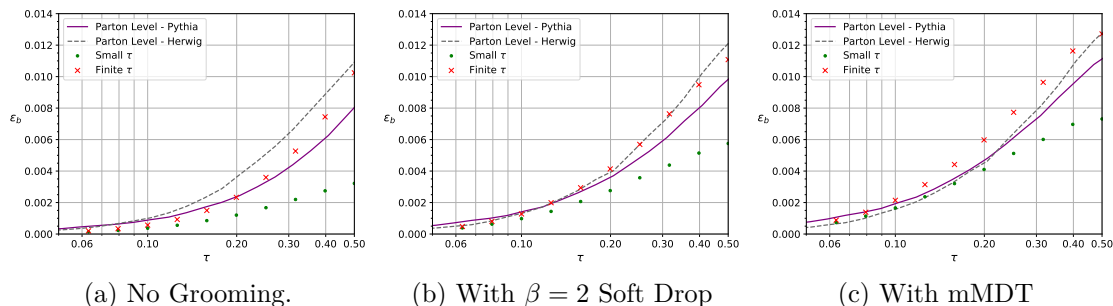


Figure 4. A comparison of our calculation for background tagging rate against parton shower Monte Carlo simulations for no grooming, grooming with Soft Drop $\beta = 2$, and grooming with mMDT.

on our calculation: no grooming, Soft Drop with $\beta = 2$ pre grooming, and pre-grooming with mMDT.

In all cases one notes that our results are in reasonable agreement with parton shower predictions, given the uncertainties of the calculations and the shower predictions due to subleading terms present in each case, and also reflected in the difference between Herwig and Pythia showers. In the case of no pre-grooming or grooming with $\beta = 2$ Soft Drop, our finite τ calculation is clearly an improvement of the small τ calculation over a wide range of τ values. Where jets are pre-groomed with mMDT, the finite τ effects we include still have a sizeable impact and improve agreement with the parton showers as $\tau \rightarrow \frac{1}{2}$, however, at smaller values of τ it is not clear that agreement with the parton showers is improved by their inclusion. This is potentially due to the fact that the leading logs in this case are single logs and we do not include any sources of next-to-leading logarithms (or their interplay with the τ dependence), other than the finite τ corrections we introduced here.

Figure 4 also shows increasing differences between results from parton showers as the level of grooming decreases. Hence the mMDT result, involving more aggressive grooming, is in better agreement between the two shower descriptions over a wider range in τ , while the un-groomed case shows the largest differences. This is likely due to the differences in the modelling of soft gluon effects between the two showers, which is ameliorated by grooming.

5 Signal jets

Here we consider the action of the Y_m -Splitter method with a τ_{32} cut on the top quark initiated signal jet. In ref. [34] studies were carried out for top jets with a range of tagging methods including Y_m -Splitter both with mMDT and Soft Drop pre-grooming. Here one has to account, in principle, for gluon radiation in both the top production and top decay processes. In the highly boosted limit the top quark is similar to a light quark and the role of soft gluon radiation and its resummation therefore becomes as important as for the background QCD case. In particular in the boosted limit one can ignore the dead-cone effect [71], which does not affect our logarithmic accuracy. We shall also consider soft gluon energies well above the top width where we can neglect additional details of the soft gluon emission pattern studied for instance in [72]. In the region relevant to our studies we can

therefore consider soft emissions as arising from a single fast moving colour charge aligned with the initial top quark direction.⁸

In spite of these simplifying dynamical assumptions, for top jets, the resummation of large logarithms for the tagging and grooming combinations we consider is more complicated than for the case of background jets. In particular the three-pronged structure of the jet can arise in multiple ways including from the electroweak decay of the top system as well as from soft gluon emission effects. Therefore as in ref. [34] our targeted accuracy will be lower for the signal case and shall omit double logarithms in ζ , ρ/ρ_{\min} and other similar ratios. We shall mainly aim at capturing leading logarithms in m^2/p_T^2 where m is a mass-scale which is at most of the order of the top mass.

5.1 Jet mass distribution for top jets

We start by computing the fraction of top jets tagged by simply requiring the invariant mass to be within some mass window. Radiation produced by the virtual top quark emerging from the hard process can be recombined with the final top decay products to form the final jet. Placing an upper limit on the jet mass therefore directly constrains this radiation and results in a Sudakov form factor precisely as for a light quark jet. We restrict ourselves to the case where the lower edge of the mass window is below the top mass, so that jets containing all of the top decay products will have mass larger than this. Of course, there will be some fraction of events where not all of the top decay products are reconstructed as a single jet, however such configurations are suppressed by a power of $\frac{m_t}{p_T}$ [29] and hence can be neglected to our accuracy. We can then write the tagged fraction of events as

$$\Sigma(\rho) = \frac{1}{\sigma_0} \int |M_{t \rightarrow bq\bar{q}}|^2 d\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \Theta_{\text{Clust}} S_{\text{QCD}}, \quad (5.1)$$

where $|M_{t \rightarrow bq\bar{q}}|^2$ is the squared matrix-element for the top decay, $d\Phi_3$ is the three-body phase-space in the collinear approximation and Θ_{Clust} is the jet clustering condition as for the background case (see eqs. (4.25) and (4.28)). The normalisation factor σ_0 is just the result without considering QCD corrections i.e. the squared matrix-element for top decay integrated over the final state phase-space with the jet clustering requirement. The factor S_{QCD} takes into account the constraint on QCD radiation through limiting the jet mass. Given that the jet mass can be expressed in terms of multiple soft gluon emissions such that $\rho = \rho_t + \sum_i \rho_i$, with $\rho_t = m_t^2/(R^2 p_T^2)$, the constraint on ρ_i just produces a Sudakov form factor which factorises from the integral over the top-decay phase-space to give:

$$\Sigma(\rho) = S_{\text{QCD}} = e^{-R(\rho - \rho_t)}, \quad (5.2)$$

where $R(\rho - \rho_t)$ is the standard jet mass Sudakov evaluated to NLL accuracy [73] at the shifted scale $\rho - \rho_t$.⁹

⁸To account for large logarithms with a τ_{32} cut we will also need to consider additional collinear radiation from the colour charges arising from W decay within the top jet.

⁹Although we use the full heavy jet mass radiator evaluated to NLL accuracy the result is only accurate to modified LL accuracy for our case. In particular we neglect non-global [74] and clustering logarithms [75] that are relevant here at NLL level.

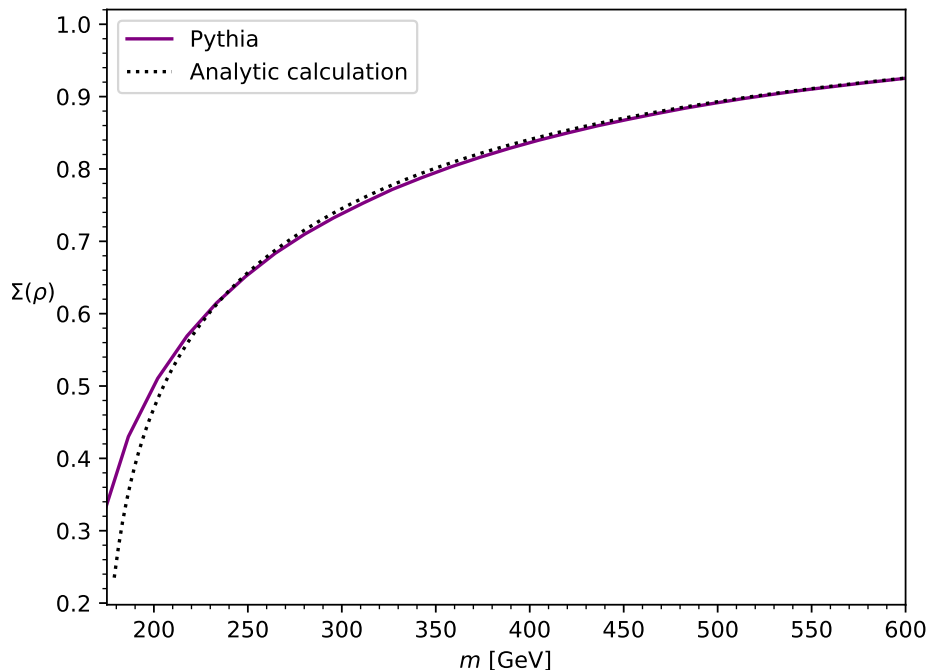


Figure 5. A comparison between our analytical calculation of the cumulative jet mass distribution $\Sigma(\rho)$ for top quark initiated jets with $\rho = m^2/R^2p_T^2$, with $p_T > 2$ TeV and the same distribution derived from Pythia simulations.

In order to test this result and the approximations inherent in deriving it, we compare our result to expectations from Pythia 8. For our Pythia 8 study we choose the lower edge of the mass window to be 10 GeV below the top mass which serves to further reject events where the top decay constituents are not recombined into the final jet. Effects contributing at the lower edge of the mass range should thus only differ from our result by numerically small effects. Pythia 8 was used to create a sample of 1 million $t\bar{t}$ events, with UE, MPI and hadronisation deactivated, and Fastjet [62] was then used to find CA jets with $R = 1$. Figure 5 shows the integrated jet mass distribution as the upper limit on the mass range is varied. Our analytical estimate is in good agreement with the distribution obtained with Pythia and Fastjet. As m approaches the top mass the agreement between our calculation and Pythia slightly worsens which is to be expected as effects which we neglect, including non-perturbative effects, become relevant for values of m very close to the top mass.

5.2 Top jets with Y_m -Splitter

Next we consider the application of Y_m -Splitter to the tagging of top jets. This was already studied in ref. [34] where it was noted that the signal case had a number of additional complications relative to the description of QCD background jets, which made the attainment of leading logarithmic accuracy in each of the parameters ρ , ρ_{\min}/ρ and ζ substantially harder. For this reason only basic leading logarithmic accuracy in ρ (or equivalently in ρ_{\min}) was targeted which allowed for a simplified treatment of the Sudakov form factor. Consistently with the accuracy goal of that article, complications including the possibility

of soft gluon emissions giving one of the three prongs found by the tagger, and the interplay with the mass window constraint were neglected. The results, broadly speaking, gave a reasonable description of the main behaviour seen with parton showers, but the agreement was not as good as seen for QCD background jets.

Here, prior to discussing N-subjettiness, we shall attempt to at least partially address some of the complications that are mentioned above for pure Y_m -Splitter. In particular we now consider in the soft-collinear limit, the situation where a single soft emission can be de-clustered as one of the prongs found by the tagger in addition to the case where the de-clustered prongs arise from the electroweak top decay process. Let us start by considering the result at leading-order i.e. neglecting all QCD radiative corrections. Then we can write

$$\Sigma(\rho, \rho_{\min}, \zeta) = \frac{1}{\sigma_0} \int |M_{t \rightarrow b\bar{q}\bar{q}}|^2 d\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \Theta_{\text{Clust}} \Theta_{Y_m\text{-Splitter}}, \quad (5.3)$$

where

$$\Theta_{Y_m\text{-Splitter}} = \Theta(\text{Min}(\rho_{12}, \rho_{13}, \rho_{23}) > \rho_{\min}) \Theta(\text{Min}(z_1, z_2, z_3) > \zeta), \quad (5.4)$$

and 1, 2, 3 refer to the three prongs identified by Y_m -Splitter. As there is no soft enhancement to the top decay matrix element, we use only a collinear approximation for the pairwise invariant masses, $\rho_{ij} = x_i x_j \theta_{ij}^2$ in our calculations of the leading-order top decay. Next we consider QCD radiative corrections in the soft and collinear limit. We first take into account the situation that no soft gluon emissions are de-clustered as a prong by Y_m -Splitter. This imposes a constraint on real emissions in addition to the constraint on jet mass, which comes from the requirement that the soft emission must set a smaller $\text{gen-}k_t$ distance than those set by the three-pronged top system. Labelling the soft emission by i we then have that $\min(d_{i1}, d_{i2}, d_{i3}) < \min(d_{12}, d_{13}, d_{23})$. This complicated constraint simplifies in the soft and strongly-ordered limit responsible for the leading double logarithms we seek. To be more precise, the three-pronged top decay results in relatively energetic particles owing to the lack of soft enhancement in the electroweak decay. For a soft gluon emission to set a comparable $\text{gen-}k_t$ distance it must be emitted at a relatively large angle compared to the opening angle between the top decay products, $1 \gg \theta_i^2 \gg \theta_{ij}^2$, where θ_i is the angle w.r.t. the jet axis or equivalently the emitting top quark. In this region we can approximate the angle made by the soft emission with any given prong from the top decay simply by the angle w.r.t. the jet axis which allows us to write the $\text{gen-}k_t$ distance for the gluon as $z_i \theta_i^2$.

In addition to the $\text{gen-}k_t$ distance, the soft emissions are also subject to the jet mass constraint as before. Therefore the argument of the Sudakov corresponds to whichever is the tighter constraint which gives

$$\begin{aligned} \Sigma^{(0)}(\rho, \rho_{\min}, \zeta) &= \frac{1}{\sigma_0} \int |M_{t \rightarrow b\bar{q}\bar{q}}|^2 d\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \\ &\quad \times \Theta_{\text{Clust}} \Theta_{Y_m\text{-Splitter}} e^{-R(\min(d_{12}, d_{13}, d_{23}, \rho - \rho_t))}, \end{aligned} \quad (5.5)$$

where by $\Sigma^{(0)}$ we mean the contribution where we enforce that no soft gluons can give one of the 3 prongs found by the tagger.

Next we correct this picture by allowing a soft emission to form one of the prongs found by Y_m -Splitter, a situation that can first arise at order α_s . Consider a single gluon emerging from the de-clustering process before one of the top decay products, and thus being identified as a prong. This gluon is constrained so that it has energy fraction $z > \zeta$ and sets a minimum pairwise mass with the other prongs (labelled 1 and 2) of m_{\min} , i.e. $\min(\rho_{1g}, \rho_{2g}) > \rho_{\min}$, where g labels the gluon. The gluon must also not set a jet mass which pushes the jet outside of the mass window. The emission of a single soft gluon factorises from the top decay process and gives an order α_s contribution to the pre-factor. Subsequent gluon emissions are constrained by the requirement of not being de-clustered as a prong as well as being subject to the jet mass constraint and again give rise to a Sudakov suppression. Hence we obtain the result:

$$\begin{aligned} \Sigma^{(1)}(\rho, \rho_{\min}, \zeta) &= \frac{1}{\sigma_0} \int |M_{t \rightarrow b\bar{q}\bar{q}}|^2 d\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \\ &\times \int dz \frac{d\theta^2}{\theta^2} \frac{\alpha_s(z\theta p_t) C_F}{\pi} p_{gq}(z) \Theta(z\theta^2 < \rho - \rho_{\text{top}}) \Theta_{\text{Clust}} \\ &\times \sum_{i < j \neq k} \left(\Theta(d_{ij} < \text{Min}(d_{ik}, d_{kj})) \Theta(z\theta^2 > d_{ij}) \Theta(\text{Min}(\rho_{k(ij)}, z\theta^2) > \rho_{\min}) \right. \\ &\times \left. \Theta(\text{Min}(z, z_k, (z_i + z_j) > \zeta) e^{-R(\text{Min}(d_{k(ij)}, z\theta^2, \rho - \rho_t - z\theta^2))} \right). \end{aligned} \quad (5.6)$$

In the above result the first line gives the pre-factor which, aside from the usual squared matrix-element and phase-space integration for top decay, now also has the QCD pre-factor coming from real emission of the soft gluon. The three prongs are given by the soft gluon, a clustered pair of particles (ij) from the top decay and the remaining particle k arising from the top decay. The condition $\Theta(z\theta^2 > d_{ij})$ alongside the requirement that $z > \zeta$ ensures that the soft gluon is de-clustered as a prong.¹⁰ The condition $\Theta(\text{Min}(\rho_{k(ij)}, z\theta^2) > \rho_{\min})$ is the ρ_{\min} condition where again we used the fact that at our accuracy we can replace the gluon angle w.r.t. a given prong by that w.r.t. the jet axis. Finally we discuss the Sudakov which has as argument $(\text{Min}(d_{k(ij)}, z\theta^2, \rho - \rho_t - z\theta^2))$, reflecting the competing constraints on subsequent soft emissions. Firstly we have that emissions must not set a gen- k_t distance larger than the smallest gen- k_t distance amongst the 3 prongs found by Y_m -Splitter, given by $\text{Min}(d_{k(ij)}, z\theta^2)$. Secondly we have that the soft emissions must not push the jet out of the mass window, i.e. the jet mass should be below ρ . Taking into account the additional soft emission we now have as a prong, this condition implies that for multiple subsequent emissions i we must have $\sum_i \rho_i < \rho - \rho_t - z\theta^2$. Taken together these conditions, on gen- k_t and mass, produce the Sudakov in eq. (5.6).

It is additionally possible for two soft emissions to be resolved i.e. form two of the prongs found by Y_m -Splitter. This occurs at order α_s^2 with only modest logarithmic enhancements¹¹ and hence such contributions are suppressed relative to the terms we include. We therefore omit them here. We also note that we have ignored soft emissions from the

¹⁰Note that here we used the same leading-logarithmic simplification for the gen- k_t distance for soft gluon emissions that led to the result in eq. (5.3).

¹¹We remind the reader that resolved emissions are constrained in several ways. They need to have energy larger than ζ as well as a mass large enough to satisfy the ρ_{\min} condition but not large enough to

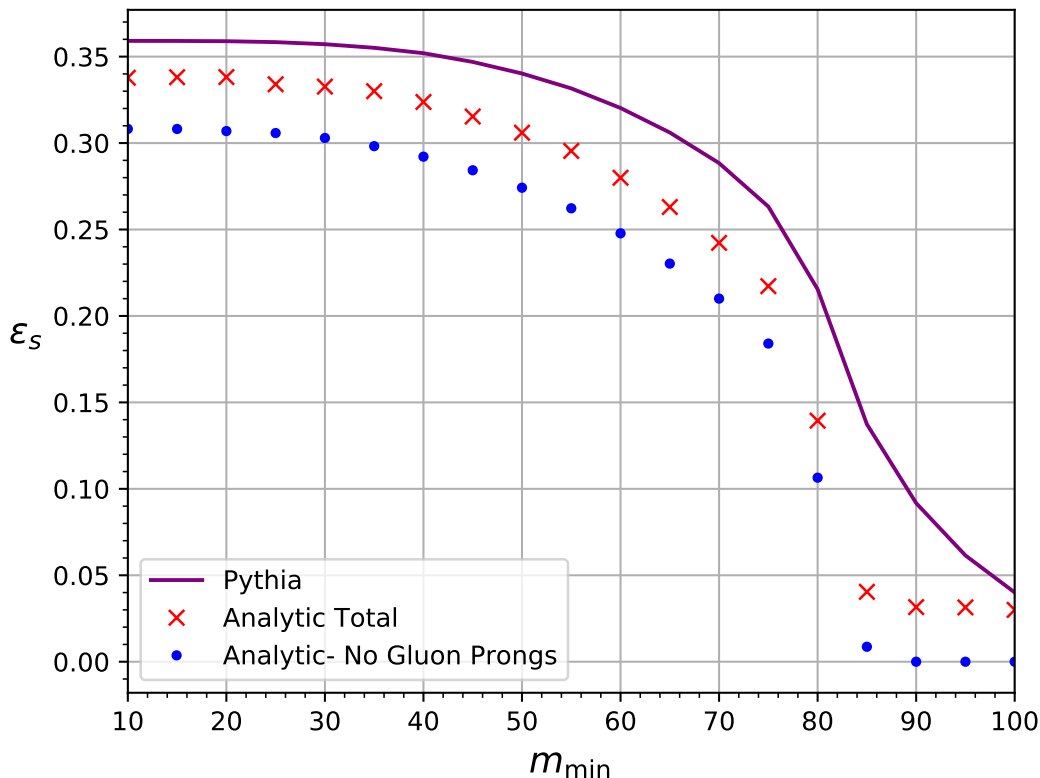


Figure 6. A comparison between different levels of approximation in analytical calculations and a Pythia simulation for top jets tagged with Y_m -Splitter in a mass range $163 \text{ GeV} < m < 225 \text{ GeV}$ with $\zeta = 0.05$ as a function of m_{\min} .

$q\bar{q}$ system produced by the splitting of the W boson. Soft emissions from this dipole are restricted in angle, by virtue of angular ordering, to have an angle less than that of the $q\bar{q}$ pair. Since they are part of the top system they also do not contribute to a shift in mass. Hence to our leading logarithmic accuracy they can also be ignored.

Our results are compared to Pythia 8 in figure 6, where we plot the signal efficiency as a function of m_{\min} (cf. similar plots in ref. [34]). We show our results for both cases with (red crosses) and without (blue dots) a resolved gluon prong. Our analytics agrees in both cases with the general behaviour seen with Pythia and we note an improved agreement with Pythia when the $\Sigma^{(1)}$ contribution, amounting to an $\mathcal{O}(15\%)$ correction, is included. As before we choose the lower limit of the mass window to be 10 GeV below the top mass.

5.3 Y_m -Splitter with grooming for signal jets

Next we examine the impact of pre-grooming with Soft Drop on our results for Y_m -Splitter applied to top jets. Relative to results from previous studies [34] here we also account for the possibility of a resolved gluon prong as in the previous subsection. The result of pre-grooming with mMDT or Soft Drop is again to essentially replace the Sudakov for the

push the jet out of the mass window. These constraints lead to the appearance of only modest logarithmic contributions.

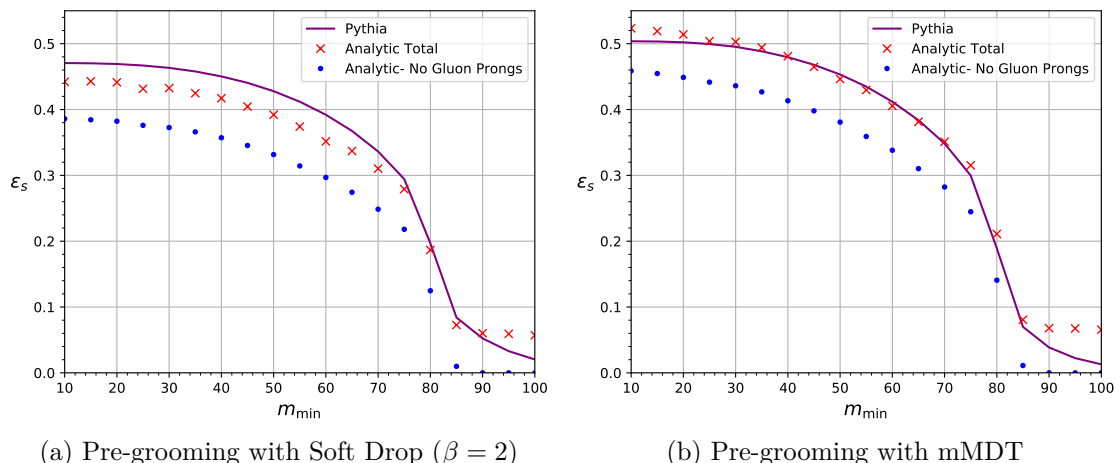


Figure 7. A comparison between our analytical calculations and a Pythia simulation for pre-groomed top jets tagged with Y_m -Splitter in a mass range $163 \text{ GeV} < m < 225 \text{ GeV}$ with $\zeta = 0.05$ as a function of m_{\min} .

un-groomed case by the Sudakov for the groomer i.e. we make the following replacements in the $\Sigma^{(0)}$ and $\Sigma^{(1)}$ terms of the un-groomed results (see eqs. (5.5) and (5.6)):

$$\begin{aligned}
 e^{-R(\min(d_{12}, d_{13}, d_{23}, \rho - \rho_t))} &\rightarrow e^{-R_{\text{mMDT/SD}}(\min(d_{12}, d_{13}, d_{23}, \rho - \rho_t))}, \\
 e^{-R(\text{Min}(d_{k(ij)}, z\theta^2, \rho - \rho_t - z\theta^2))} &\rightarrow e^{-R_{\text{mMDT/SD}}(\text{Min}(d_{k(ij)}, z\theta^2, \rho - \rho_t - z\theta^2))},
 \end{aligned}
 \tag{5.7}$$

where the suffix mMDT or SD is used to indicate the grooming variant. We note that unlike the case of the QCD background jets, we have not included R_{angle} terms in the signal case. Although such terms would in principle be present, the angular scales involved are of the order of the opening angles between top decay products. At such angular scales the radiation pattern becomes more complicated as one also needs to account for radiation from the $q\bar{q}$ dipole produced by the colour singlet W decay. Given that the terms produced are logarithms in the ratio of two small scales, i.e. of the same level of significance as $\ln \rho/\rho_{\min}$ terms, they are beyond the accuracy we aim for in the case of signal jets.

The tagged signal fraction, with our usual choice of parameter values, is compared to a Pythia simulation in figure 7, again showing the results with and without a resolved gluon prong and for grooming with SD (left) and mMDT(right). We see that except for the extreme region, where the tagged signal fraction is very small, our analytic results, especially after inclusion of the resolved gluon case, are in good overall agreement with the behaviour seen with Pythia.

5.4 Y_m -Splitter with τ_{32} and grooming for signal jets

We now wish to understand the effect of adding a cut on τ_{32} to the tagged signal distribution after application of Y_m -Splitter. We shall first consider the un-groomed case and then include the effects of grooming. We begin with the configuration where all three of the LO top decay products are identified as prongs by Y_m -Splitter. With no additional emissions τ_3 vanishes and hence a cut on τ_{32} has no impact. Adding a set of soft and collinear emissions, one has to consider how these emissions are constrained by the τ cut, the mass-window cut and the requirement that they should not give a resolved prong on applying Y_m -Splitter.

We first introduce an approximation into our definition of τ_2 which is valid to within the overall accuracy we can obtain with our current calculations for the signal case, i.e. LL accuracy in ρ with neglect of logs in ratios of mass scales and ζ . Consider the region of phase space where say $d_{12} < \min(d_{13}, d_{23})$, so that the first de-clustering will lead to two gen- k_t axes lying along p_3 and $p_1 + p_2$. In this region of phase space, to leading order where there are no additional emissions, $\tau_2 = z_1 \theta_{1,12}^2 + z_2 \theta_{2,12}^2$. As the $p_1 + p_2$ direction will be aligned more with the harder of partons 1 and 2 we make the approximation that the gen- k_t axis is aligned with this parton, so that to LO we can approximate $\tau_2 = \min(d_{12}, d_{13}, d_{23})$. As there is no logarithmic enhancement associated with the leading order decay of the top, this approximation will introduce an $\mathcal{O}(1)$ rescaling of the argument of the Sudakov factor, which is consistent with an NLL correction and hence beyond our LL accuracy.

When considering the role of additional soft emissions let us first consider, as in section 5.3 before, primary emissions at a large angle to the opening angles of the top decay system. Regardless of which of the gen- k_t axes these emissions are closer to, their contribution to τ_3 and τ_2 may always be approximated by $\sum_i \rho_i$, where $\rho_i = z_i \theta_i^2$ and θ_i is the emission angle w.r.t. the emitting top quark direction. The constraint on emissions due to the τ_{32} cut is then $\tau_{32} \approx \frac{\sum_i \rho_i}{\min(d_{12}, d_{13}, d_{23}) + \sum_i \rho_i} < \tau$ which gives the constraint $\sum_i \rho_i < \min(d_{12}, d_{13}, d_{23}) \frac{\tau}{1-\tau}$. For $\tau < 1/2$ this subjettiness constraint overcomes the constraint from Y_m -Splitter, $\rho_i < \min(d_{12}, d_{13}, d_{23})$ and hence the argument of the primary emission Sudakov depends only on the competing subjettiness and jet mass constraints.

Until now we have neglected the role of secondary radiation from the $q\bar{q}$ system (arising from W decay) since these emissions bring only enhancements in ratios of similar mass scales. If we wish to obtain a good description of the signal with a τ cut including also the region where $\tau \ll 1$, we need to consider all sources of double-logarithmic corrections in τ . Secondary emissions are a source of such double-logarithmic terms and hence we include them here. The secondary emission terms are given by taking into account soft and collinear emissions from the q and \bar{q} with the constraint that the emission angle is smaller than $\theta_{q\bar{q}}$ the opening angle of the $q\bar{q}$ dipole. This leads to results which have the same form as the corresponding results for the background case (see eq. (4.6)) with z_a replaced by z_q and θ_a by $\theta_{q\bar{q}}$ for emission from q and similarly for emission from the \bar{q} . We note that secondary emissions are part of the decaying top system and hence do not contribute to a shift in mass so that the jet mass constraint is irrelevant here.

Thus we can write

$$\Sigma^{\tau < \frac{1}{2}}(\rho_{\min}, \zeta, \tau) = \frac{1}{\sigma_0} \int |M_{t \rightarrow b q \bar{q}}|^2 d\Phi_3 \delta\left(\frac{s_{123}}{R^2 p_t^2} - \rho_t\right) \Theta_{\text{Clust}} \Theta_{Y_m\text{-Splitter}} \frac{e^{-R - \gamma_E R'}}{\Gamma[1 + R']}, \quad (5.8)$$

with

$$\begin{aligned} R \equiv & R \left(\min \left(\frac{\tau}{1-\tau} \min(d_{12}, d_{13}, d_{23}), \rho_{\max} - \rho_t \right) \right) \\ & + R^{\text{secondary}} \left(\frac{\tau}{1-\tau} \min(d_{12}, d_{13}, d_{23}), z_q, \theta_{q\bar{q}}^2 \right) \\ & + R^{\text{secondary}} \left(\frac{\tau}{1-\tau} \min(d_{12}, d_{13}, d_{23}), z_{\bar{q}}, \theta_{q\bar{q}}^2 \right), \end{aligned} \quad (5.9)$$

where ρ_{\max} is the upper limit on the jet mass. Finally we account for the effect of grooming. To take this into account one makes the usual replacement of the primary emission radiator

by its groomed counterpart. An additional subtlety that is present here is the existence of R_{angle} terms (see eq. (4.22)) which originate from emissions which are not visible to the groomer as they are shielded by larger angle emissions that stop the grooming. Such terms have been ignored for the signal since they are complicated to account for and produce only logarithms of mass ratios which we neglect. However in the presence of a τ cut such terms also induce double logarithms in τ as described by eq. (4.23). A consistent description of the double logs in τ should also include the double logarithm originating here while we can neglect all other details associated to this term. Grooming is therefore included through the replacement of the radiator as

$$R\left(\min\left(\frac{\tau}{1-\tau}\min(d_{12}, d_{13}, d_{23}), \rho_{\text{max}} - \rho_t\right)\right) \rightarrow R_{\text{mMDT}}\left(\min\left(\frac{\tau}{1-\tau}\min(d_{12}, d_{13}, d_{23}), \rho_{\text{max}} - \rho_t\right)\right) + R_{\text{angle}}(\tau), \quad (5.10)$$

where, at fixed coupling, $R_{\text{angle}}(\tau) = \frac{C_F \alpha_s}{2\pi} \ln^2 \tau$.

We have thus far not considered the case where a soft gluon is resolved as a Y_m -Splitter prong, which we took into account in the previous subsections. For such a configuration, the effect of the τ cut is actually to constrain the phase space of partons arising from the LO top decay. As the electroweak top decay is not logarithmically enhanced, the restriction from the τ cut leads to a suppression proportional to τ . Given that the configuration with a resolved gluon prong is already suppressed by a power of α_s a further suppression with τ implies that we may ignore this term while still retaining a reasonable description of the overall behaviour.¹²

Equation (5.8) is evaluated and compared to the same distribution derived from simulations using Pythia in figure 8. Although given the accuracy of the shower and the analytic calculations (each of which is leading-logarithmic albeit with inclusion of some key NLL effects), one would expect to see the moderate level of difference that can be observed in the figure, it is noticeable that the behaviour in τ is well captured by the analytics especially for the un-groomed case and for pre-grooming with Soft Drop. For grooming with mMDT there is good agreement at smaller τ and a deviation at larger values of τ . Here, given that the leading logarithms are single logarithms, the analytics and the shower would each only contain (at best) a correct leading-logarithmic description, but with potentially larger differences from spurious NLL effects in the shower and their interplay with τ . Moreover our neglect of configurations where a gluon is one of the resolved prongs from Y_m -Splitter would also lead to differences at larger values of τ where the power suppression with τ , which was a factor in our neglecting this configuration, will be less pronounced. Neglect of such configurations may have more of an impact on the distributions where jets are pre-groomed, as they can allow the jet to be tagged even if one of the electroweak top decay products is groomed away.

¹²We remind the reader that the value of τ that gives the highest signal significance is $\tau \sim 0.2$. We have checked numerically, by studying specific configurations, that the power suppression with τ holds at leading order.

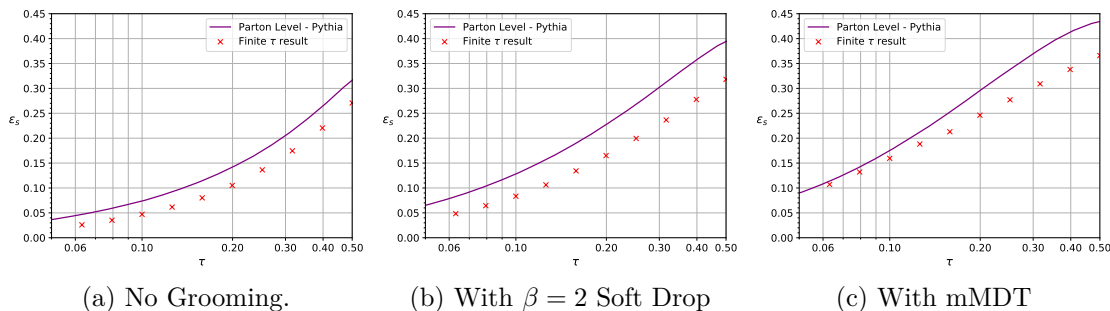


Figure 8. Comparison between our analytic calculation (crosses) and Pythia for the tagged signal distribution as a function of the cut on τ_{32} for jets without pre-grooming (left), with pre-grooming using Soft Drop (centre) and with pre-grooming using the mMDT.

We note that eq. (5.8) for the signal case reflects a few features that are different to the corresponding results for the QCD background. In particular for signal jets there is a lack of soft and collinear enhancements in the pre-factor resulting in the absence of the Hypergeometric function. Also, to our accuracy, the jet mass constraint does not affect the distribution for small enough τ cuts, or large enough ρ_{\max} , as a result of the fixed invariant mass of the leading-order system. This is clear from the argument of the Sudakov factor in equation (5.8) which contains a competition between the τ cut and the mass-window. For a given τ_{32} cut we can estimate the threshold below which m_{\max} should be taken if varying it is to have an effect on the tagged signal fraction:

$$m_{\max}^2 < m_t^2 + p_T^2 \frac{\tau}{1 - \tau} \min(d_{12}, d_{13}, d_{23}). \quad (5.11)$$

For top jets, where $\min(d_{12}, d_{13}, d_{23})p_T^2$ may be roughly approximated by the W boson mass squared, we estimate that, for $\tau = 0.3$ and $m_t = 173$ GeV, the jet mass constraint will not significantly affect the signal efficiency unless $m_{\max} \lesssim 181$ GeV. In reality there will not be a hard threshold but some range of parameters over which the Sudakov suppression transitions from being due to the cut on τ_{32} to being due to the jet mass constraint. The application of this will be discussed further in the next section.

6 Exploiting jet mass cuts

In this section we discuss a notable feature of our calculations in terms of the differences between signal and background jets. As suggested by eq. (5.11), one can reduce the cut on jet mass m_{\max} , without impacting the signal until we reach a critical value depending on τ . Until we reach this point reducing m_{\max} results in a decrease in the background tagging rate and hence an increase in performance. While our analytic studies are somewhat simplified and in particular neglect subleading terms, it is interesting to study the extent to which our observations may apply to parton shower studies when subleading effects are present. Figure 9 shows, using both analytic calculations (left) and parton level MC simulations (right), how the signal tagging rate varies with m_{\max} for several fixed τ cuts both without grooming and with grooming via Soft Drop and the mMDT.

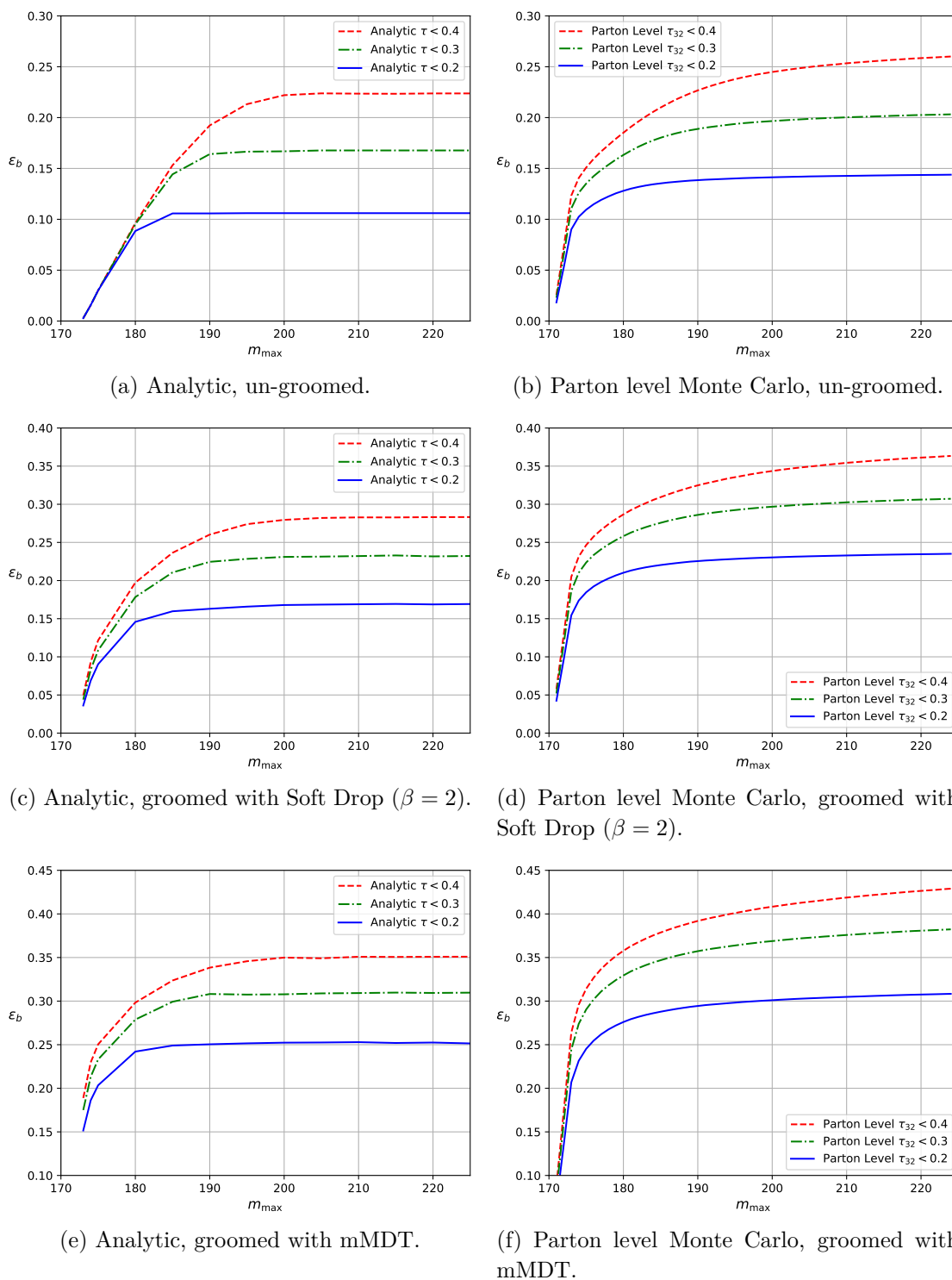


Figure 9. Analytic and Monte Carlo parton level curves showing how the signal tagging rate varies with m_{\max} .

For the signal distribution the overall shape and dependence on τ is well described by our calculation, although, as before, there is some difference in the overall normalisation. The difference between our calculation and the distribution derived from MC worsens for smaller values of m_{\max} , which should be expected, as non-perturbative effects, which can not be completely removed from parton shower simulations, will start to play more of a role in this region. While the signal tagging rate derived from MC simulations does not flatten off to the same extent as the analytic calculations do as m_{\max} is increased, it is clear that beyond a certain value of m_{\max} the signal efficiency depends only very weakly on m_{\max} .

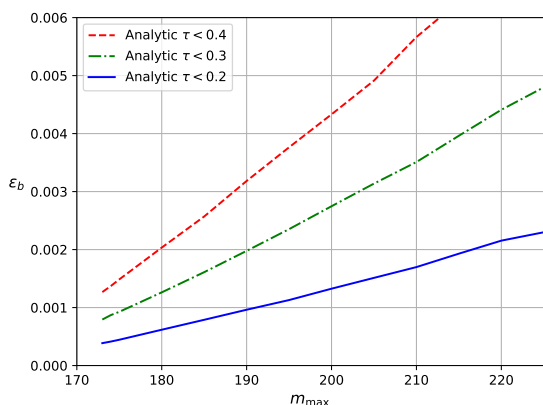
Figure 10 shows similar plots for the case of quark jets. Our analytic predictions are again seen to be in overall good agreement with the Pythia shower capturing the m_{\max} and τ dependences. It is notable that the jet mass constraint affects the background tagging rate in the same way for any cut on τ_{32} , as there are not two competing scales in the Sudakov factor. This opens up the possibility to improve the performance of the tagging procedure by reducing m_{\max} so that the signal tag rate remains approximately constant whilst removing a significant portion of the background.

One may wonder, given the effectiveness of a tight cut on the jet mass, what improvement is gained by cutting on τ_{32} in these circumstances. Figure 11 also shows a curve generated by varying m_{\max} over the range 173 GeV to 225 GeV, but with no cut on τ_{32} . In this case the signal significance is higher than cutting on τ_{32} with $m_{\max} = 225$ GeV, but cutting on τ_{32} with $m_{\max} = 180$ GeV is still the highest performing tagging procedure.

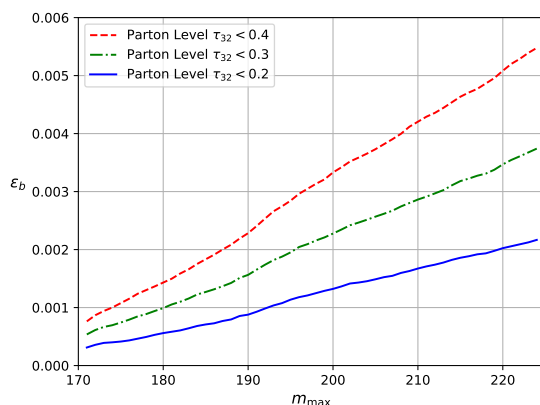
We now investigate the impact of non-perturbative corrections on this tagging procedure as m_{\max} is varied. Figure 12 shows the resilience [76] to non-perturbative effects, defined as $\zeta_{\text{NP}} = \left(\frac{\Delta\epsilon_s^2}{\langle\epsilon_s\rangle^2} + \frac{\Delta\epsilon_b^2}{\langle\epsilon_b\rangle^2} \right)^{-1/2}$ where $\Delta\epsilon$ is the difference between the parton and hadron level tagging efficiency and $\langle\epsilon\rangle$ is the mean of the two, for jets pre-groomed with mMDT, as m_{\max} is varied, for three different values of τ . To construct the resilience 10 million $q\bar{q}$ events and 1 million $t\bar{t}$ events were generated at both parton and hadron level using Pythia. From figure 12 we see that the resilience to non-perturbative effects does not strongly depend on m_{\max} in the range considered, even with m_{\max} as low as 180 GeV. By contrast, reducing the cut on τ_{32} from 0.4 to 0.2 results in a marked drop in resilience. It would therefore be beneficial, in terms of reducing the impact of non-perturbative effects, to take τ not too small, say $\tau = 0.4$, while imposing a rather tight cut on the jet mass to provide the discriminating power. These cuts provide a signal significance of around 6 with a signal efficiency around 0.35. This is both a higher signal efficiency and significance than was reported in section 3 with $m_{\max} = 225$ GeV and $\tau = 0.2$, the highest significance achieved with the higher value of m_{\max} .

7 Conclusions

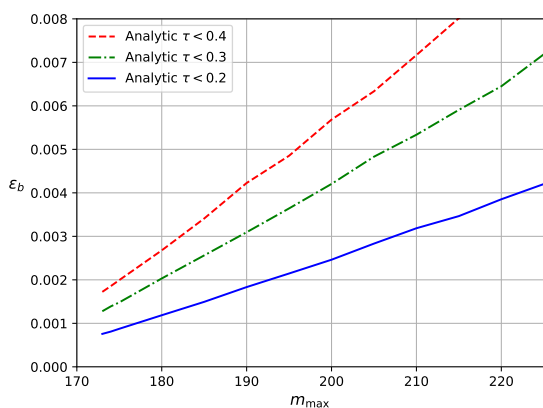
In this article we have studied top-tagging from first principles of QCD, as part of a larger program to understand the features of tagging and grooming methods in a model-independent fashion. We chose a combination of methods, starting from the application of a prong-finding step aimed at tagging three-pronged decays and rejecting background,



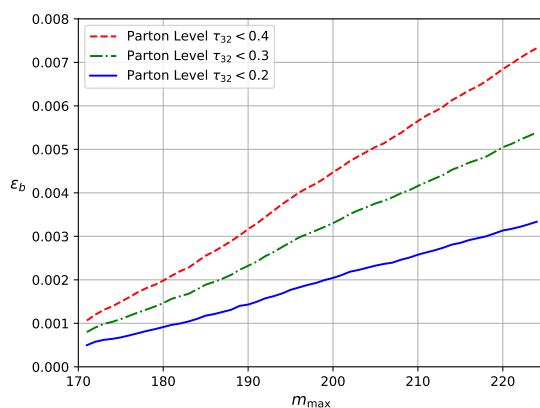
(a) Analytic, un-groomed.



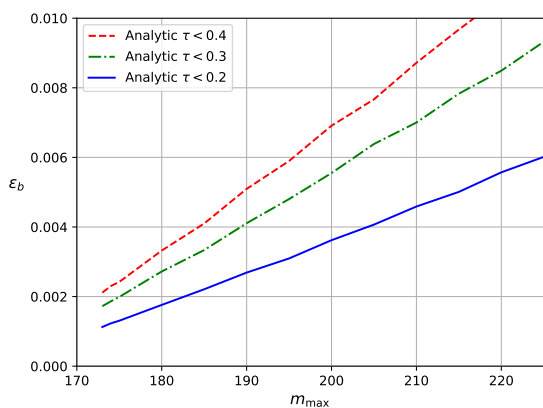
(b) Parton level Monte Carlo, un-groomed.



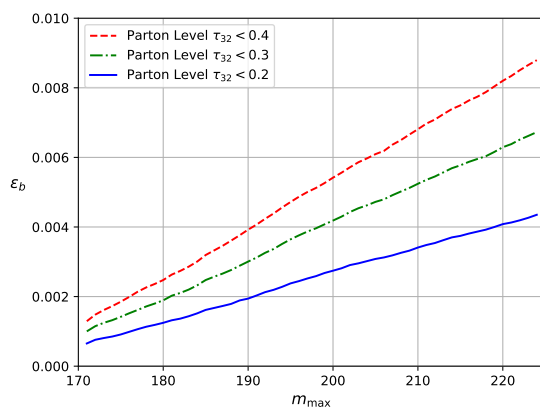
(c) Analytic, groomed with Soft Drop ($\beta = 2$).



(d) Parton level Monte Carlo, groomed with Soft Drop ($\beta = 2$).



(e) Analytic, groomed with mMDT.



(f) Parton level Monte Carlo, groomed with mMDT.

Figure 10. Analytic and Monte Carlo (parton level) curves showing how the background tagging rate varies with m_{\max} .

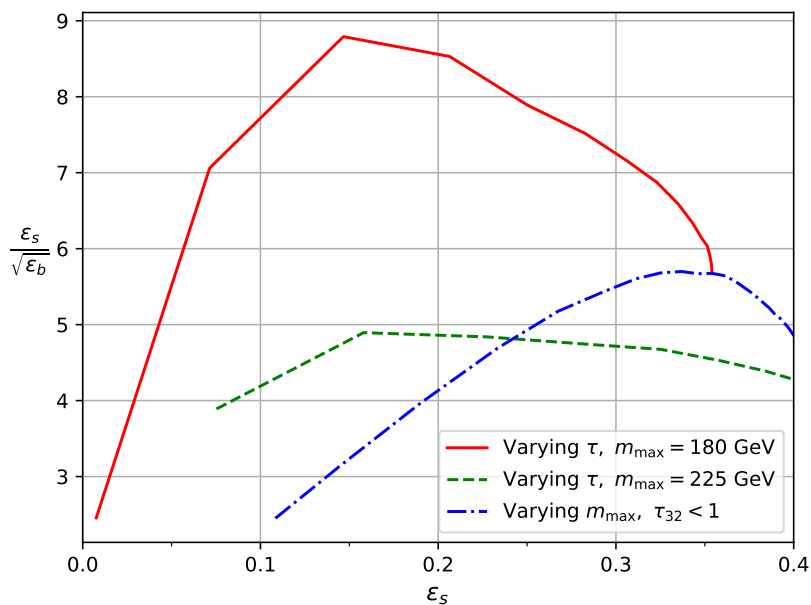


Figure 11. Signal significance against efficiency for three variations on the tagging procedure. All jets are groomed with mMDT and tagged with Y_m -Splitter. Either τ or m_{\max} is varied with a fixed cut placed on the other. The samples were produced using Pythia with hadronisation and UE activated.

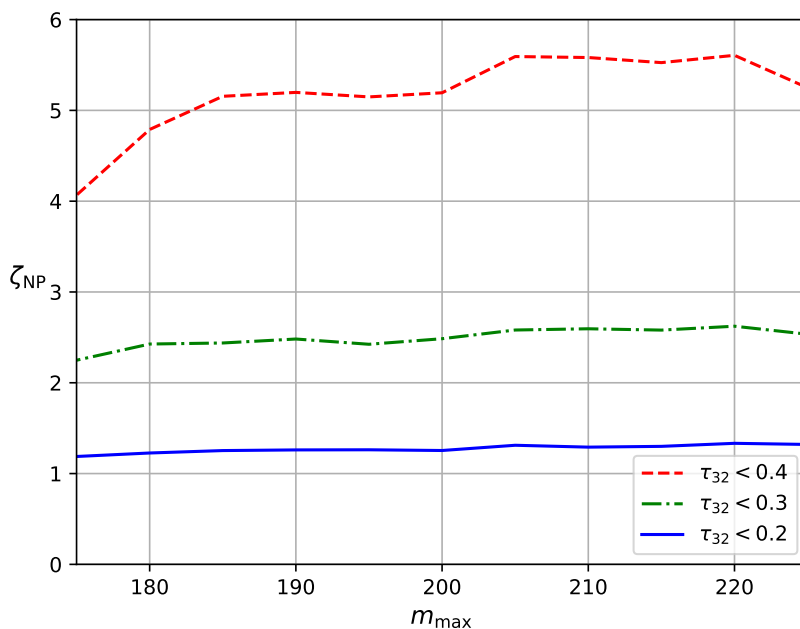


Figure 12. A measure of resilience to non perturbative effects as m_{\max} varies for three cuts on τ_{32} . Jets are pre-groomed with mMDT.

followed by a radiation constraining step. We also pre-groom jets with both the mMDT and with Soft Drop with $\beta = 2$ to reduce non-perturbative contributions. For prong finding we have used Y_m -Splitter, an adaptation of Y-splitter introduced in ref. [30] while as a radiation constraining shape variable we have applied the N-subjettiness ratio τ_{32} with $\beta = 2$. While our specific choices (use of Y_m -Splitter for prong-finding and $\beta = 2$ for τ_{32}) are helpful in somewhat simplifying analytical studies, combinations similar to the ones used here have commonly been employed, including for experimental studies involving top-tagging [54–56].

We started by carrying out Monte Carlo studies which provided some of the motivation for what followed in terms of yielding information on performance, resilience to non-perturbative effects, and optimal parameter choices for our combination of methods. Next we turned to studying QCD background jets. Here we have built on previous work [34] on understanding top-tagging, and in particular Y_m -Splitter, by including the constraint from τ_{32} . We have derived results for the double differential distribution in jet mass and τ_{32} as well as for the cumulant where we integrate over τ_{32} with the condition $\tau_{32} < \tau$. We obtained a result in the limit of small τ and then included finite τ corrections along similar lines to the studies in ref. [35]. We also performed studies both with and without pre-grooming with mMDT and Soft Drop ($\beta = 2$). We compared our results to those from both Herwig and Pythia showers and in all cases we saw that our analytical calculations are able to capture the essential impact of the tagging, shape-variable and grooming steps.

We then turned to studying signal jets. We found that in the highly boosted limit a mass window constraint gives rise to a simple Sudakov form factor which is in good agreement with Pythia results. We then added Y_m -Splitter as in ref. [34], but improved upon previous calculations by also considering a situation, at order α_s , where a soft gluon can be one of the prongs resolved by Y_m -Splitter. Including this contribution we found the results to be in significantly better agreement with Pythia than was the case with previous results where such a correction was not considered [34]. We then studied the impact of a τ cut, including finite τ effects and also considering pre-grooming with mMDT and Soft Drop $\beta = 2$. In all cases, in spite of the complexity of the problem, our simplifying approximations were sufficient to capture, the basic behaviour, i.e the τ dependence over a wide range in τ , seen also with parton showers. Remaining differences with parton showers were at a level that was consistent with our expectations from missing subleading terms. One immediately exploitable outcome of our analytic results was the suggestion that using a tighter mass cut than our default choice (for a given τ) would reduce the background rather than the signal while not significantly affecting the resilience to non-perturbative effects. This finding was used to show how a highly performant and resilient method could be developed using our combination of tools.

Finally we would say that although the combination of methods we have considered gives rise to a highly non-trivial observable, we have demonstrated that analytical methods can still give substantial insight into the basic physics mechanisms that control the performance of such tool combinations. Further systematic improvements on the results we have obtained are possible, with the inclusion of subleading logarithmic terms being one avenue to pursue. Also, while our specific choice of tools is helpful for analytical studies

and was taken mainly for convenience, combinations similar to the ones here have been in widespread use, and have not been analytically understood thus far. We believe that our studies should therefore encourage analytical investigations of other similar combinations, including for example variants using $\tau_{32}^{(\beta=1)}$, and help develop a more complete picture of the distinct role played by the different elements and/or steps that form part of a number of top tagging methods.

Acknowledgments

We would like to thank Gregory Soyez and Kiran Ostrolenk for helpful discussions. This work has been funded by the U.K. Science and Technologies Facilities Council under grant ST/P000800/1. JH thanks the U.K. Science and Technologies Facilities Council (STFC) for a PhD studentship award.

A Finite τ approximations at fixed order

Here we explain the impact of the different approximations on emissions ρ_a and ρ_b , which enter the pre-factor for the resummed expressions in the small τ limit and for finite τ . In order to illustrate this we examine the differential distribution in a fixed-coupling approximation and at order $\bar{\alpha}^3$, which is the first order at which τ_{32} is non-zero. The finite τ result truncated at order $\bar{\alpha}^3$ is (from eq. (4.12)):

$$\frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} \stackrel{\tau \ll 1/2}{\simeq} \bar{\alpha}^2 \frac{1}{1-\tau} \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho-\rho_a} \Theta\left(\rho_a > \frac{1-\tau}{2-\tau}\rho\right) \Theta_{\rho_{\min}} R'((\rho-\rho_a)\tau), \quad (\text{A.1})$$

where in a fixed-coupling approximation, and to leading-logarithmic accuracy, we have,

$$R'((\rho-\rho_a)\tau) = \bar{\alpha} \left(\ln \frac{1}{\rho-\rho_a} + \ln \frac{1}{\tau} \right). \quad (\text{A.2})$$

We remind the reader that

$$\Theta_{\rho_{\min}} = \Theta\left(\min\left((\rho-\rho_a)(1-\tau), z_a z_b \max\left(\frac{\rho_a}{z_a}, \frac{(\rho-\rho_a)(1-\tau)}{z_b}\right)\right) > \rho_{\min}\right), \quad (\text{A.3})$$

where $\frac{\rho_a}{z_a} = \theta_a^2$ and $\frac{(\rho-\rho_a)(1-\tau)}{z_b} = \frac{\rho_b}{z_b} = \theta_b^2$. For our illustrative purposes, let us take a specific contribution to the pre-factor that arises from the region $\theta_a^2 \gg \theta_b^2$ and $\rho_b = (\rho-\rho_a)(1-\tau) < z_b \rho_a$. We also assume values of the parameters so that $\frac{\rho_{\min}}{\rho} > \zeta$, again purely as part of our illustrative example.¹³ The conclusions we derive will apply to other regions of phase-space and parameter ranges too. For the region of phase space considered we obtain the contribution

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} \stackrel{\tau \ll 1/2}{\simeq} \bar{\alpha}^2 \frac{1}{1-\tau} \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho-\rho_a} \Theta\left(\rho_a > \frac{1-\tau}{2-\tau}\rho\right) R'((\rho-\rho_a)\tau) \\ \times \Theta((\rho-\rho_a)(1-\tau) > \rho_{\min}) \Theta\left(\frac{\rho_a}{z_a} > \frac{(\rho-\rho_a)(1-\tau)}{z_b}\right) \\ \times \Theta((\rho-\rho_a)(1-\tau) < z_b \rho_a). \end{aligned} \quad (\text{A.4})$$

¹³Strictly speaking the condition is $\frac{\rho_{\min}}{\rho} > \frac{\zeta}{1+\zeta}$.

The directly computed small τ limit result, corresponding to the result in section 4.2.1, is

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} &= \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho} \frac{d\rho_b}{\rho_b} \Theta(\rho_b > \rho_{\min}) \Theta(\rho_b < z_b \rho) \\ &\quad \times \Theta\left(\rho > \frac{z_a}{z_b} \rho_b\right) \bar{\alpha} \left(\ln \frac{1}{\tau} + \ln \frac{1}{\rho_b} \right) \end{aligned} \quad (\text{A.5})$$

which gives

$$\frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} = \frac{1}{2} \bar{\alpha}^2 \ln \frac{1}{\zeta} \ln^2 \frac{\rho}{\rho_{\min}} \times \bar{\alpha} \left(\ln \frac{1}{\rho_{\min}} + \ln \frac{1}{\tau} + \frac{1}{3} \ln \frac{\rho_{\min}}{\rho} \right). \quad (\text{A.6})$$

Evaluating the small τ limit of the exact result in our chosen configuration i.e. eq. (A.4), gives:

$$\frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} = \frac{1}{2} \bar{\alpha}^2 \ln \frac{1}{\zeta} \ln^2 \frac{\rho}{\rho_{\min}} \times \bar{\alpha} \left(\ln \frac{1}{\rho_{\min}} + \ln \frac{1}{\tau} + \frac{1}{3} \ln \frac{\rho_{\min}}{\rho} \right) + \frac{3}{4} \bar{\alpha}^3 \zeta(3) \ln \frac{1}{\zeta}, \quad (\text{A.7})$$

which differs from eq. (A.6) by a highly subleading $\bar{\alpha}^3 \zeta(3) \ln \frac{1}{\zeta}$ term. The full τ dependent result from eq. (A.4) reads

$$\begin{aligned} \frac{\rho\tau}{\sigma} \frac{d^2\sigma}{d\rho d\tau} &= \frac{\bar{\alpha}^3}{1-\tau} \left[\frac{1}{2} \ln \frac{1}{\zeta} \ln^2 \frac{\rho}{\rho_{\min}} \times \left(\ln \frac{1}{\rho_{\min}} + \ln \frac{1}{\tau} + \frac{1}{3} \ln \frac{\rho_{\min}}{\rho} + \ln(1-\tau) \right) \right. \\ &\quad \left. - \ln \frac{1}{\zeta} \text{Li}_3 \left(\frac{1}{\tau-1} \right) \right]. \end{aligned} \quad (\text{A.8})$$

We note that retaining the finite τ effects results in the appearance of three features: firstly there is the overall $1/(1-\tau)$ multiplicative term, which has a significant impact on the result beyond the small τ region and is important to retain. Secondly there is a $\ln(1-\tau)$ term in addition to the large logarithms we resum. Given that we do not resum logarithms of $1-\tau$, and indeed focus on the region $\tau \sim 0.2$, this constitutes a negligible contribution relative to the logarithms we resum, dominated by the $\ln \frac{1}{\rho_{\min}}$ term. Finally there is a highly subleading $\bar{\alpha}^3 \ln \frac{1}{\zeta}$ term accompanied by a trilogarithm in $1-\tau$ which we can safely neglect. Hence ignoring the τ dependence in the pre-factor, other than the $1/(1-\tau)$ term, is a valid approximation for our work.

B Direct calculation of the cumulative background distribution

Here we provide an alternate derivation of $\rho \frac{d\Sigma(\tau)}{d\rho}$ that directly derives this distribution as opposed to integrating the double differential. We can start from the standard factorised formula for any number of emissions, similar to eq. (4.9) but instead of fixing τ_{32} we set an upper bound $\tau_{32} < \tau$:

$$\begin{aligned} \frac{\rho}{\sigma} \frac{d\sigma}{d\rho d\tau} &= \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_0^1 \frac{d\rho_a}{\rho_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_0^{\rho_a} \frac{d\rho_b}{\rho_b} \Theta\left(\min\left(\rho_b, z_a z_b \max\left(\frac{\rho_a}{z_a}, \frac{\rho_b}{z_b}\right)\right) > \rho_{\min} \right) \\ &\quad \times \exp\left[- \int_0^1 R(\rho') \frac{d\rho'}{\rho'} \right] \sum_{p=1}^{\infty} \frac{1}{p!} \prod_{i=1}^p \int_0^{\rho_b} R'(\rho_i) \frac{d\rho_i}{\rho_i} \rho \delta\left(\rho - \rho_a - \rho_b - \sum_{i \neq a,b} \rho_i \right) \\ &\quad \times \Theta((\rho - \rho_a)(1-\tau) < \rho_b). \end{aligned} \quad (\text{B.1})$$

The delta function can now be used to do the sum over emissions labelled with i , where it is crucial to notice that for any i , $\rho_i < \rho - \rho_a - \rho_b$ as implied by the delta function, and that the upper limit on ρ_i of ρ_b is weaker than this for $\tau < \frac{1}{2}$. Using the standard jet mass result for the sum over emissions, as we did before, we arrive at

$$\begin{aligned} \frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \stackrel{\tau < 1/2}{=} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_{\frac{1-\tau}{2}\rho}^{\rho} \frac{d\rho_a}{\rho_a} \int_{(1-\tau)(\rho-\rho_a)}^{\min(\rho-\rho_a, \rho_b)} \frac{d\rho_b}{\rho_b} \frac{\rho}{\rho - \rho_a - \rho_b} \\ \times \Theta_{\rho_{\min}}(\rho_a, \rho_b, \tau, \rho_{\min}, z_a, z_b) R'(\rho - \rho_a - \rho_b) \\ \times \frac{\exp[-R(\rho - \rho_a - \rho_b) - \gamma_E R'(\rho - \rho_a - \rho_b)]}{\Gamma[1 + R'(\rho - \rho_a - \rho_b)]}. \end{aligned} \quad (\text{B.2})$$

We now wish to integrate over ρ_b which we note contains two regions, $\rho_a < \rho - \rho_a$ and $\rho_a > \rho - \rho_a$, the former of which vanishes if we neglect the τ dependence of the ρ_a integral limits as in section 4.3.2.¹⁴ Enforcing the condition $\rho_b > \rho_{\min}$, which is embodied in $\Theta_{\rho_{\min}}$ gives an upper limit on ρ_a of $\rho - \frac{\rho_{\min}}{1-\tau}$, which, within our accuracy we can approximate as ρ . To carry out this integral within single logarithmic accuracy, we can expand the radiator about some fixed ρ_b which we take as $(\rho - \rho_a)(1 - \tau_0)$ so that:

$$R(\rho - \rho_a - \rho_b) \simeq R((\rho - \rho_a)\tau_0) - R'((\rho - \rho_a)\tau_0) \ln\left(\frac{\rho - \rho_a - \rho_b}{(\rho - \rho_a)\tau_0}\right) + \mathcal{O}(R''), \quad (\text{B.3})$$

where τ_0 should be chosen close to τ as values of ρ_b close to $(\rho - \rho_a)(1 - \tau)$ are expected to dominate the integral. The integral can then be carried out to give

$$\begin{aligned} \frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \stackrel{\tau < 1/2}{=} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_{\frac{1}{2}\rho}^{\rho} \frac{d\rho_a}{\rho_a} \frac{\rho}{\rho - \rho_a} \Theta_{\rho_{\min}}(\rho_a, \rho_{\min}, z_a, z_b) \\ \times \left(\frac{\tau}{\tau_0}\right)^{R'((\rho-\rho_a)\tau_0)} {}_2F_1\left(1, R'(\rho - \rho_a)\tau_0, 1 + R'(\rho - \rho_a)\tau_0, \tau\right) \\ \times \frac{\exp[-R((\rho - \rho_a)\tau_0) - \gamma_E R'((\rho - \rho_a)\tau_0)]}{\Gamma[1 + R'((\rho - \rho_a)\tau_0)]}, \end{aligned} \quad (\text{B.4})$$

in perfect agreement with eq. (4.16).

Although less convenient for making contact with the result reported in eq. (4.16), we could equally well have integrated over ρ_a , leaving the ρ_b integral to be done numerically, as we could have done in section 4.3.1. To do this one would expand $R(\rho - \rho_a - \rho_b)$ around $\rho_a = \rho - \frac{\rho_b}{1-\tau}$ which would lead to:

$$\begin{aligned} \frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \stackrel{\tau < 1/2}{=} \bar{\alpha}^2 \int_{\zeta}^1 \frac{dz_a}{z_a} \int_{\zeta}^1 \frac{dz_b}{z_b} \int_{\rho_{\min}}^{\frac{\rho}{2}} \frac{d\rho_b}{\rho_b} \frac{\rho}{\rho - \rho_b} \left(\frac{\tau(1-\tau_0)}{(1-\tau)\tau_0}\right)^{R'(\rho_b \frac{\tau_0}{1-\tau_0})} \Theta_{\rho_{\min}}(\rho_b, \rho_{\min}, z_a, z_b) \\ \times {}_2F_1\left(1, R'\left(\rho_b \frac{\tau_0}{1-\tau_0}\right), 1 + R'\left(\rho_b \frac{\tau_0}{1-\tau_0}\right), \frac{\rho_b \tau_0}{(1-\tau_0)(\rho - \rho_b)}\right) \\ \times \frac{\exp[-R(\rho_b \frac{\tau_0}{1-\tau_0}) - \gamma_E R'(\rho_b \frac{\tau_0}{1-\tau_0})]}{\Gamma[1 + R'(\rho_b \frac{\tau_0}{1-\tau_0})]}, \end{aligned} \quad (\text{B.5})$$

where again, τ_0 should be taken close to τ , and any τ dependence in the leading order pre-factor has been neglected.

¹⁴The neglected term is proportional to $\alpha^3 L^2 L_\rho$ where L_ρ is a log of ρ or ρ_{\min} while L is a log of the ratio or ζ . This is clearly beyond our accuracy.

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

References

- [1] M.H. Seymour, *Searches for new particles using cone and cluster jet algorithms: a comparative study*, *Z. Phys. C* **62** (1994) 127 [[INSPIRE](#)].
- [2] J.M. Butterworth, B.E. Cox and J.R. Forshaw, *WW scattering at the CERN LHC*, *Phys. Rev. D* **65** (2002) 096014 [[hep-ph/0201098](#)] [[INSPIRE](#)].
- [3] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [[arXiv:0802.2470](#)] [[INSPIRE](#)].
- [4] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination algorithms and jet substructure: pruning as a tool for heavy particle searches*, *Phys. Rev. D* **81** (2010) 094023 [[arXiv:0912.0033](#)] [[INSPIRE](#)].
- [5] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, *Phys. Rev. D* **80** (2009) 051501 [[arXiv:0903.5081](#)] [[INSPIRE](#)].
- [6] D. Krohn, J. Thaler and L.-T. Wang, *Jet trimming*, *JHEP* **02** (2010) 084 [[arXiv:0912.1342](#)] [[INSPIRE](#)].
- [7] A. Abdesselam et al., *Boosted objects: a probe of beyond the standard model physics*, *Eur. Phys. J. C* **71** (2011) 1661 [[arXiv:1012.5412](#)] [[INSPIRE](#)].
- [8] A. Altheimer et al., *Jet substructure at the Tevatron and LHC: new results, new tools, new benchmarks*, *J. Phys. G* **39** (2012) 063001 [[arXiv:1201.0008](#)] [[INSPIRE](#)].
- [9] A. Altheimer et al., *Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd–27th of July 2012*, *Eur. Phys. J. C* **74** (2014) 2792 [[arXiv:1311.2708](#)] [[INSPIRE](#)].
- [10] D. Adams et al., *Towards an understanding of the correlations in jet substructure*, *Eur. Phys. J. C* **75** (2015) 409 [[arXiv:1504.00679](#)] [[INSPIRE](#)].
- [11] A.J. Larkoski, I. Moult and B. Nachman, *Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning*, *Phys. Rept.* **841** (2020) 1 [[arXiv:1709.04464](#)] [[INSPIRE](#)].
- [12] D.E. Kaplan, K. Rehermann, M.D. Schwartz and B. Tweedie, *Top tagging: a method for identifying boosted hadronically decaying top quarks*, *Phys. Rev. Lett.* **101** (2008) 142001 [[arXiv:0806.0848](#)] [[INSPIRE](#)].
- [13] T. Plehn, G.P. Salam and M. Spannowsky, *Fat jets for a light Higgs*, *Phys. Rev. Lett.* **104** (2010) 111801 [[arXiv:0910.5472](#)] [[INSPIRE](#)].
- [14] CMS collaboration, *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging*, *CMS-PAS-JME-09-001* (2009).
- [15] CMS collaboration, *Boosted top jet tagging at CMS*, *CMS-PAS-JME-13-007* (2014).
- [16] G. Brooijmans, *High p_T hadronic top quark identification. Part I: jet mass and YSplitter*, *ATL-PHYS-CONF-2008-008* (2008).
- [17] J. Thaler and L.-T. Wang, *Strategies to identify boosted tops*, *JHEP* **07** (2008) 092 [[arXiv:0806.0023](#)] [[INSPIRE](#)].

- [18] P. Gras et al., *Systematics of quark/gluon tagging*, *JHEP* **07** (2017) 091 [[arXiv:1704.03878](#)] [[INSPIRE](#)].
- [19] C. Frye, A.J. Larkoski, J. Thaler and K. Zhou, *Casimir meets Poisson: improved quark/gluon discrimination with counting observables*, *JHEP* **09** (2017) 083 [[arXiv:1704.06266](#)] [[INSPIRE](#)].
- [20] A.J. Larkoski and E.M. Metodiev, *A theory of quark vs. gluon discrimination*, *JHEP* **10** (2019) 014 [[arXiv:1906.01639](#)] [[INSPIRE](#)].
- [21] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, Springer, Germany (2019) [[arXiv:1901.10342](#)] [[INSPIRE](#)].
- [22] R. Kogler et al., *Jet substructure at the Large Hadron Collider: experimental review*, *Rev. Mod. Phys.* **91** (2019) 045003 [[arXiv:1803.06991](#)] [[INSPIRE](#)].
- [23] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [[arXiv:1307.0007](#)] [[INSPIRE](#)].
- [24] M. Dasgupta, A. Fregoso, S. Marzani and A. Powling, *Jet substructure with analytical methods*, *Eur. Phys. J. C* **73** (2013) 2623 [[arXiv:1307.0013](#)] [[INSPIRE](#)].
- [25] A.J. Larkoski, G.P. Salam and J. Thaler, *Energy correlation functions for jet substructure*, *JHEP* **06** (2013) 108 [[arXiv:1305.0007](#)] [[INSPIRE](#)].
- [26] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft drop*, *JHEP* **05** (2014) 146 [[arXiv:1402.2657](#)] [[INSPIRE](#)].
- [27] A.J. Larkoski, I. Moult and D. Neill, *Analytic boosted boson discrimination*, *JHEP* **05** (2016) 117 [[arXiv:1507.03018](#)] [[INSPIRE](#)].
- [28] M. Dasgupta, L. Schunk and G. Soyez, *Jet shapes for boosted jet two-prong decays from first-principles*, *JHEP* **04** (2016) 166 [[arXiv:1512.00516](#)] [[INSPIRE](#)].
- [29] M. Dasgupta, A. Powling and A. Siodmok, *On jet substructure methods for signal jets*, *JHEP* **08** (2015) 079 [[arXiv:1503.01088](#)] [[INSPIRE](#)].
- [30] M. Dasgupta, A. Powling, L. Schunk and G. Soyez, *Improved jet substructure methods: Y-splitter and variants with grooming*, *JHEP* **12** (2016) 079 [[arXiv:1609.07149](#)] [[INSPIRE](#)].
- [31] G.P. Salam, L. Schunk and G. Soyez, *Dichroic subjetteness ratios to distinguish colour flows in boosted boson tagging*, *JHEP* **03** (2017) 022 [[arXiv:1612.03917](#)] [[INSPIRE](#)].
- [32] A.J. Larkoski, I. Moult and D. Neill, *Analytic boosted boson discrimination at the Large Hadron Collider*, [arXiv:1708.06760](#) [[INSPIRE](#)].
- [33] A.J. Larkoski, I. Moult and D. Neill, *Factorization and resummation for groomed multi-prong jet shapes*, *JHEP* **02** (2018) 144 [[arXiv:1710.00014](#)] [[INSPIRE](#)].
- [34] M. Dasgupta, M. Guzzi, J. Rawling and G. Soyez, *Top tagging: an analytical perspective*, *JHEP* **09** (2018) 170 [[arXiv:1807.04767](#)] [[INSPIRE](#)].
- [35] D. Napoletano and G. Soyez, *Computing N-subjettiness for boosted jets*, *JHEP* **12** (2018) 031 [[arXiv:1809.04602](#)] [[INSPIRE](#)].
- [36] F.A. Dreyer, G.P. Salam and G. Soyez, *The Lund jet plane*, *JHEP* **12** (2018) 064 [[arXiv:1807.04758](#)] [[INSPIRE](#)].
- [37] F.A. Dreyer and H. Qu, *Jet tagging in the Lund plane with graph networks*, *JHEP* **03** (2021) 052 [[arXiv:2012.08526](#)] [[INSPIRE](#)].

- [38] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images — Deep learning edition*, *JHEP* **07** (2016) 069 [[arXiv:1511.05190](#)] [[INSPIRE](#)].
- [39] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet substructure classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **93** (2016) 094034 [[arXiv:1603.09349](#)] [[INSPIRE](#)].
- [40] J. Barnard, E.N. Dawe, M.J. Dolan and N. Rajcic, *Parton shower uncertainties in jet substructure analyses with deep neural networks*, *Phys. Rev. D* **95** (2017) 014018 [[arXiv:1609.00607](#)] [[INSPIRE](#)].
- [41] P.T. Komiske, E.M. Metodiev and J. Thaler, *Energy flow polynomials: a complete linear basis for jet substructure*, *JHEP* **04** (2018) 013 [[arXiv:1712.07124](#)] [[INSPIRE](#)].
- [42] G. Kasieczka, T. Plehn, M. Russell and T. Schell, *Deep-learning top taggers or the end of QCD?*, *JHEP* **05** (2017) 006 [[arXiv:1701.08784](#)] [[INSPIRE](#)].
- [43] A. Butter, G. Kasieczka, T. Plehn and M. Russell, *Deep-learned top tagging with a Lorentz layer*, *SciPost Phys.* **5** (2018) 028 [[arXiv:1707.08966](#)] [[INSPIRE](#)].
- [44] S. Macaluso and D. Shih, *Pulling out all the tops with computer vision and deep learning*, *JHEP* **10** (2018) 121 [[arXiv:1803.00107](#)] [[INSPIRE](#)].
- [45] E.A. Moreno et al., *JEDI-net: a jet identification algorithm based on interaction networks*, *Eur. Phys. J. C* **80** (2020) 58 [[arXiv:1908.05318](#)] [[INSPIRE](#)].
- [46] H. Qu and L. Gouskos, *ParticleNet: jet tagging via particle clouds*, *Phys. Rev. D* **101** (2020) 056019 [[arXiv:1902.08570](#)] [[INSPIRE](#)].
- [47] S.H. Lim and M.M. Nojiri, *Morphology for jet classification*, [arXiv:2010.13469](#) [[INSPIRE](#)].
- [48] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni and G.P. Salam, *Logarithmic accuracy of parton showers: a fixed-order study*, *JHEP* **09** (2018) 033 [Erratum *ibid.* **03** (2020) 083] [[arXiv:1805.09327](#)] [[INSPIRE](#)].
- [49] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [[arXiv:2002.11114](#)] [[INSPIRE](#)].
- [50] D.E. Soper and M. Spannowsky, *Finding top quarks with shower deconstruction*, *Phys. Rev. D* **87** (2013) 054012 [[arXiv:1211.3140](#)] [[INSPIRE](#)].
- [51] L.G. Almeida, S.J. Lee, G. Perez, G. Sterman and I. Sung, *Template overlap method for massive jets*, *Phys. Rev. D* **82** (2010) 054034 [[arXiv:1006.2035](#)] [[INSPIRE](#)].
- [52] J. Thaler and K. Van Tilburg, *Identifying boosted objects with N -subjettiness*, *JHEP* **03** (2011) 015 [[arXiv:1011.2268](#)] [[INSPIRE](#)].
- [53] A.J. Larkoski, I. Moult and D. Neill, *Building a better boosted top tagger*, *Phys. Rev. D* **91** (2015) 034035 [[arXiv:1411.0665](#)] [[INSPIRE](#)].
- [54] ATLAS collaboration, *Search for $W' \rightarrow tb \rightarrow qqbb$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **75** (2015) 165 [[arXiv:1408.0886](#)] [[INSPIRE](#)].
- [55] ATLAS collaboration, *Search for charged Higgs bosons in the $H^\pm \rightarrow tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *JHEP* **03** (2016) 127 [[arXiv:1512.03704](#)] [[INSPIRE](#)].
- [56] CMS collaboration, *Top tagging with new approaches*, [CMS-PAS-JME-15-002](#) (2016).
- [57] ATLAS collaboration, *Reconstruction of high mass $t\bar{t}$ resonances in the lepton + jets channel*, [ATL-COM-PHYS-2009-255](#) (2009).

- [58] G. Brooijmans, *High p_T hadronic top quark identification*, [ATL-PHYS-CONF-2008-008](#) (2008).
- [59] CMS collaboration, *Top tagging with new approaches*, [CMS-PAS-JME-15-002](#) (2016).
- [60] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279, 4, 1998 [[hep-ph/9907280](#)] [[INSPIRE](#)].
- [61] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, [JHEP 08 \(1997\) 001](#) [[hep-ph/9707323](#)] [[INSPIRE](#)].
- [62] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, [Eur. Phys. J. C 72 \(2012\) 1896](#) [[arXiv:1111.6097](#)] [[INSPIRE](#)].
- [63] J. Thaler and K. Van Tilburg, *Maximizing boosted top identification by minimizing N -subjettiness*, [JHEP 02 \(2012\) 093](#) [[arXiv:1108.2701](#)] [[INSPIRE](#)].
- [64] A.J. Larkoski and J. Thaler, *Unsafe but calculable: ratios of angularities in perturbative QCD*, [JHEP 09 \(2013\) 137](#) [[arXiv:1307.1699](#)] [[INSPIRE](#)].
- [65] A.J. Larkoski, S. Marzani and J. Thaler, *Sudakov safety in perturbative QCD*, [Phys. Rev. D 91 \(2015\) 111501](#) [[arXiv:1502.01719](#)] [[INSPIRE](#)].
- [66] ATLAS collaboration, *Performance of top-quark and W -boson tagging with ATLAS in Run 2 of the LHC*, [Eur. Phys. J. C 79 \(2019\) 375](#) [[arXiv:1808.07858](#)] [[INSPIRE](#)].
- [67] S. Catani, L. Trentadue, G. Turnock and B. Webber, *Resummation of large logarithms in e^+e^- event shape distributions*, [Nucl. Phys. B 407 \(1993\) 3](#).
- [68] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, [Comput. Phys. Commun. 168 \(2005\) 78](#) [[hep-ph/0404043](#)] [[INSPIRE](#)].
- [69] Wolfram Research Inc., *Mathematica, Version 11.3*.
- [70] M. Bahr et al., *HERWIG++ physics and manual*, [Eur. Phys. J. C 58 \(2008\) 639](#) [[arXiv:0803.0883](#)] [[INSPIRE](#)].
- [71] Y.L. Dokshitzer, V.A. Khoze and S.I. Troian, *On specific QCD properties of heavy quark fragmentation ('dead cone')*, [J. Phys. G 17 \(1991\) 1602](#) [[INSPIRE](#)].
- [72] L.H. Orr, Y.L. Dokshitzer, V.A. Khoze and W.J. Stirling, *Gluon radiation and top width effects*, [hep-ph/9307338](#) [[INSPIRE](#)].
- [73] S. Catani, G. Turnock and B. Webber, *Heavy jet mass distribution in e^+e^- annihilation*, [Phys. Lett. B 272 \(1991\) 368](#).
- [74] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*, [Phys. Lett. B 512 \(2001\) 323](#) [[hep-ph/0104277](#)] [[INSPIRE](#)].
- [75] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa and S. Marzani, *Non-global logarithms and jet algorithms in high- p_T jet shapes*, [JHEP 08 \(2010\) 064](#) [[arXiv:1004.3483](#)] [[INSPIRE](#)].
- [76] J. Bendavid et al., *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*, [arXiv:1803.07977](#).